(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 April 2003 (17.04.2003)

PCT

(10) International Publication Number
WO 03/031573 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US02/31460

(22) International Filing Date: 3 October 2002 (03.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/327,520 5 October 2001 (05.10.2001) US

(71) Applicant *(for all designated States except US)*: **COLEY PHARMACEUTICAL GMBH** [DE/DE]; Elisabeth-Selbert-Strasse 9, 40764 Langenfeld (DE).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **LIPFORD, Grayson** [US/DE]; Benrode Str. 86, 40597 Duesseldorf (DE).

(74) Agent: **STEELE, Alan, W.**; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).

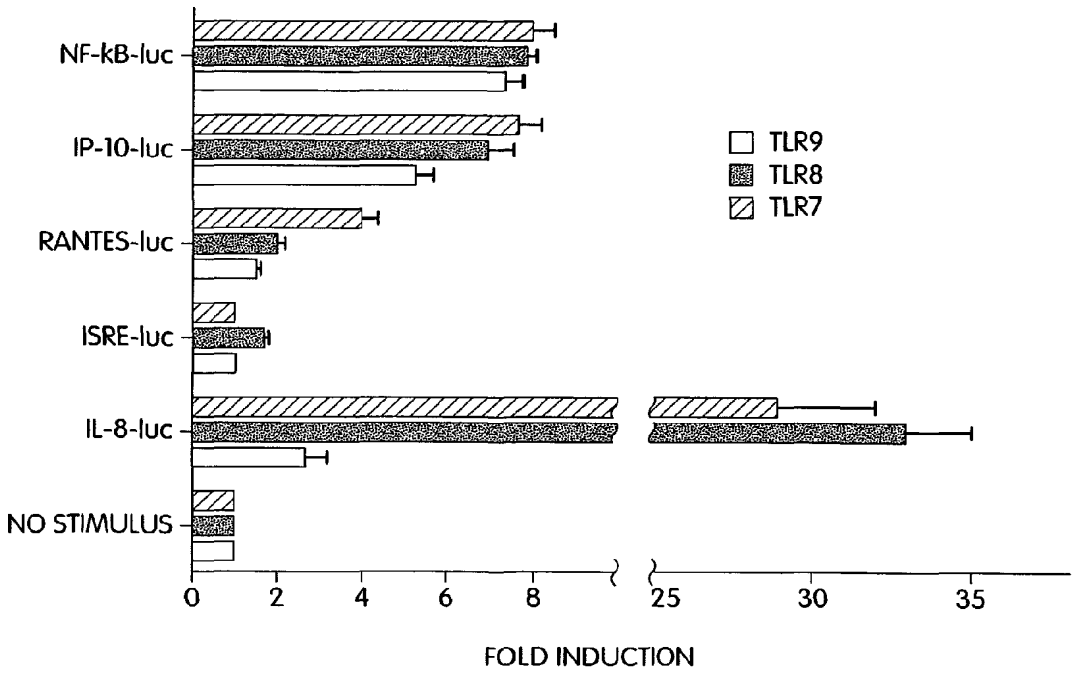
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *without international search report and to be republished upon receipt of that report*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: TOLL-LIKE RECEPTOR 3 SIGNALING AGONISTS AND ANTAGONISTS




(57) Abstract: Compositions and methods are provided to identify, characterize, and optimize immunostimulatory compounds, their agonists and antagonists, working through TLR3.

WO 03/031573 A2

# TOLL-LIKE RECEPTOR 3 SIGNALING AGONISTS AND ANTAGONISTS

## Field of the Invention

The invention pertains to signal transduction by Toll-like receptor 3 (TLR3), which is believed to be involved in innate immunity. More specifically, the invention pertains to screening methods useful for the identification and characterization of TLR3 ligands, TLR3 signaling agonists, and TLR3 signaling antagonists.

## Background of the Invention

Toll-like receptors (TLRs) are a family of at least ten highly conserved receptor proteins (TLR1 – TLR10) which recognize pathogen-associated molecular patterns (PAMPs) and act as key elements in innate immunity. As members of the pro-inflammatory interleukin-1 receptor (IL-1R) family, TLRs share homologies in their cytoplasmic domains called Toll/IL-1R homology (TIR) domains. PCT published applications PCT/US98/08979 and PCT/US01/16766. Intracellular signaling mechanisms mediated by TIRs appear generally similar, with MyD88 (Wesche H et al. (1997) *Immunity* 7:837-47; Medzhitov R et al. (1998) *Mol Cell* 2:253-8; Adachi O et al. (1998) *Immunity* 9:143-50; Kawai T et al. (1999) *Immunity* 11:115-22) and tumor necrosis factor receptor-associated factor 6 (TRAF6; Cao Z et al. (1996) *Nature* 383:443-6; Lomaga MA et al. (1999) *Genes Dev* 13:1015-24) believed to have critical roles. Signal transduction between MyD88 and TRAF6 is known to involve members of the serine-threonine kinase IL-1 receptor-associated kinase (IRAK) family, including at least IRAK-1 and IRAK-2. Muzio M et al. (1997) *Science* 278:1612-5.

Ligands for many but not all of the TLRs have been described. For instance, it has been reported that TLR2 signals in response to peptidoglycan and lipopeptides. Yoshimura A et al. (1999) *J Immunol* 163:1-5; Brightbill HD et al. (1999) *Science* 285:732-6; Aliprantis AO et al. (1999) *Science* 285:736-9; Takeuchi O et al. (1999) *Immunity* 11:443-51; Underhill DM et al. (1999) *Nature* 401:811-5. TLR4 has been reported to signal in response to lipopolysaccharide (LPS). Hoshino K et al. (1999) *J Immunol* 162:3749-52; Poltorak A et al. (1998) *Science* 282:2085-8; Medzhitov R et al. (1997) *Nature* 388:394-7. Bacterial flagellin has been reported to be a natural ligand for TLR5. Hayashi F et al. (2001) *Nature* 410:1099-1103. TLR6, in conjunction with with TLR2, has been reported to signal in response to proteoglycan. Ozinsky A et al.

(2000) *PNAS USA* 97:13766-71; Takeuchi O et al. (2001) *Int Immunol* 13:933-40. Recently it was recently reported that TLR9 is a receptor for CpG DNA. Hemmi H et al. (2000) *Nature* 408:740-5.

## Summary of the Invention

The invention provides screening methods and compositions useful for the identification and characterization of compounds which themselves signal through Toll-like receptor 3 (TLR3) or which influence signaling through TLR3. Compounds which themselves signal through TLR3 are presumptively immunostimulatory. Compounds which influence signaling through TLR3 include both agonists and antagonists of TLR3 signaling activity. The methods provided by the invention are adaptable to high throughput screening, thus accelerating the identification and characterization of previously unknown inducers, agonists, and antagonists of TLR3 signaling activity.

The methods of the invention rely at least in part on the ability to assess TLR3 signaling activity. It has surprisingly been discovered according to the present invention that reporter constructs having reporter genes under control of certain promoter response elements sensitive to TLR3 signaling activity are useful in the screening assays of the invention. For example it has been surprisingly discovered according to the present invention that a reporter gene under control of interferon-specific response element (ISRE) is sensitive to TLR3 signaling activity.

It has also surprisingly been discovered according to the present invention that screening assays for TLR ligands and other assays involving TLR signaling activity can benefit from optimization for at least one of the variables of (a) concentration of test and/or reference compound, (b) kinetics of the assay, and (c) selection of reporter. Interpretation of assay data can be influenced by each of these variables.

In one aspect the invention provides a screening method for identifying an immunostimulatory compound. The method according to this aspect of the invention involves the steps of (a) contacting a functional TLR3 with a test compound under conditions which, in absence of the test compound, permit a negative control response mediated by a TLR3 signal transduction pathway; (b) detecting a test response mediated by the TLR3 signal transduction pathway; and (c) determining the test

compound is an immunostimulatory compound when the test response exceeds the negative control response. In this and in all aspects of the invention, in one embodiment the screening method is performed on a plurality of test compounds. A test compound according to this and all aspects of the invention is in one embodiment a member of a library of compounds, preferably a combinatorial library of compounds. Also in this and in all aspects of the invention, a test compound is preferably a small molecule, a nucleic acid, a polypeptide, an oligopeptide, or a lipid. In more preferred embodiments, the test compound is a small molecule or a nucleic acid. In one embodiment a test compound that is a nucleic acid is a CpG nucleic acid.

In another aspect the invention provides a screening method for identifying an immunostimulatory compound. The method according to this aspect of the invention involves the steps of (a) contacting a functional TLR3 with a test compound under conditions which, in presence of a reference immunostimulatory compound, permit a reference response mediated by a TLR3 signal transduction pathway; (b) detecting a test response mediated by the TLR3 signal transduction pathway; and (c) determining the test compound is an immunostimulatory compound when the test response equals or exceeds the reference response. In this and other aspects of the invention, a reference immunostimulatory compound is preferably a small molecule, a nucleic acid, a polypeptide, an oligopeptide, or a lipid. In one embodiment the reference immunostimulatory compound is a CpG nucleic acid.

In a further aspect the invention provides a screening method for identifying a compound that modulates TLR3 signaling activity. The method according to this aspect of the invention involves the steps of (a) contacting a functional TLR3 with a test compound and a reference immunostimulatory compound under conditions which, in presence of the reference immunostimulatory compound alone, permit a reference response mediated by a TLR3 signal transduction pathway; (b) detecting a test-reference response mediated by the TLR3 signal transduction pathway; (c) determining the test compound is an agonist of TLR3 signaling activity when the test-reference response exceeds the reference response; and (d) determining the test compound is an antagonist of TLR3 signaling activity when the reference response exceeds the test-reference response.

In yet another aspect the invention provides a screening method for identifying species specificity of an immunostimulatory compound. The method according to this aspect of the invention involves the steps of (a) measuring a first species-specific response mediated by a TLR3 signal transduction pathway when a functional TLR3 of a first species is contacted with a test compound; (b) measuring a second species-specific response mediated by the TLR3 signal transduction pathway when a functional TLR3 of a second species is contacted with the test compound; and (c) comparing the first species-specific response with the second species-specific response. In a preferred embodiment the functional TLR3 of the first species is a human TLR3. In one preferred embodiment the functional TLR3 of the first species is a human TLR3 and the functional TLR3 of the second species is a mouse TLR3.

In preferred embodiments of the foregoing aspects of the invention, the response mediated by the TLR3 signal transduction pathway is measured quantitatively.

Also in preferred embodiments of the foregoing aspects of the invention, the functional TLR3 is expressed in a cell. For example, in one embodiment the cell is an isolated mammalian cell that naturally expresses the functional TLR3. Alternatively, in another embodiment the cell is an isolated mammalian cell that does not naturally express the functional TLR3, wherein the cell has an expression vector for TLR3. For example, in one preferred embodiment the cell is a human 293 fibroblast. In other embodiments, the functional TLR3 is part of a cell-free system.

Particularly useful in embodiments of the invention involving cells which express functional TLR3 are cells which include a reporter construct sensitive to TLR3 signaling. In one embodiment the cell includes an expression vector having an isolated nucleic acid which encodes a reporter construct selected from the group of nuclear factor-kappa B-luciferase (NF-κB-luc), IFN-specific response element-luciferase (ISRE-luc), interleukin-6-luciferase (IL-6-luc), interleukin 8-luciferase (IL-8-luc), interleukin 12 p40 subunit-luciferase (IL-12 p40-luc), interleukin 12 p40 subunit-beta galactosidase (IL-12 p40-β-Gal), activator protein 1-luciferase (AP1-luc), interferon alpha-luciferase (IFN-α-luc), interferon beta-luciferase (IFN-β-luc), RANTES-luciferase (RANTES-luc), tumor necrosis factor-luciferase (TNF-luc), IP-10-luciferase

(IP-10-luc), and interferon-inducible T cell alpha chemoattractant-luciferase (I-TAC-luc). In a preferred embodiment the reporter construct is ISRE-luc.

In one embodiment according to each of the foregoing aspects of the invention, the functional TLR3 is part of a complex with a non-TLR protein selected from the group consisting of MyD88, IL-1 receptor associated kinase 1-3 (IRAK1, IRAK2, IRAK3), tumor necrosis factor receptor-associated factor 1-6 (TRAF1 - TRAF6), IκB, NF-κB, MyD88-adapter-like (Mal), Toll-interleukin 1 receptor (TIR) domain-containing adapter protein (TIRAP), Tollip, Rac, and functional homologues and derivatives thereof. In a related embodiment functional TLR3 is part of a complex with a non-TLR protein listed above, excluding MyD88.

Also according to each of the foregoing aspects of the invention, in one embodiment the response mediated by a TLR3 signal transduction pathway is induction of a reporter gene under control of a promoter response element selected from the group consisting of ISRE, IL-6, IL-8, IL-12 p40, IFN-α, IFN-β, IFN-ω, RANTES, TNF, IP-10, and I-TAC. For example, in a preferred embodiment the reporter gene under control of a promoter response element is selected from the group consisting of ISRE-luc, IL-6-luc, IL-8-luc, IL-12 p40-luc, IL-12 p40-β-Gal, IFN-α-luc, IFN-β-luc, RANTES-luc, TNF-luc, IP-10-luc, and I-TAC-luc. In one preferred embodiment the reporter gene under control of a promoter response element is ISRE-luc. In yet another preferred embodiment the reporter gene is selected from the group consisting of IFN-α1-luc and IFN-α4-luc.

In yet another embodiment according to each of the foregoing aspects of the invention, the response mediated by a TLR3 signal transduction pathway is selected from the group consisting of (a) induction of a reporter gene under control of a minimal promoter responsive to a transcription factor selected from the group consisting of AP1, NF-κB, ATF2, IRF3, and IRF7; (b) secretion of a chemokine; and (c) secretion of a cytokine. For example, in one preferred embodiment the response mediated by a TLR3 signal transduction pathway is induction of a reporter gene selected from the group consisting of AP1-luc and NF-κB-luc. In another preferred embodiment the response mediated by a TLR3 signal transduction pathway is secretion of a type 1 IFN. In yet another preferred embodiment the response mediated by a TLR3 signal transduction

pathway is secretion of a chemokine selected from the group consisting of CCL5 (RANTES), CXCL9 (Mig), CXCL10 (IP-10), and CXCL11 (I-TAC).

The sensitivity and interpretation of the screening methods of the present invention can be optimized. Such optimization involves proper selection of any one or combination of (a) concentration of test and/or reference compound, (b) kinetics of the assay, and (c) reporter. Thus, further according to each of the first three aspects of the invention, in one embodiment the contacting a functional TLR3 with a test compound further entails, for each test compound, contacting with the test compound at each of a plurality of concentrations. For example, each test compound may be evaluated at various concentrations which differ by log increments. Also according to each of the foregoing aspects of the invention, in one embodiment the detecting is performed 4-12 hours, preferably 6-8 hours, following the contacting. Similarly, in yet another embodiment according to each of the foregoing aspects of the invention, the detecting is performed 16-24 hours following the contacting. Detecting performed 4-12 hours, preferably 6-8 hours, following the contacting is believed to be more sensitive to affinity of interaction than is detecting at later times. Detecting performed 16-24 hours or later following the contacting is believed to be more sensitive to stability and duration of receptor/ligand interaction. Furthermore, because certain reporter constructs are more sensitive to certain TLRs than others, proper matching of reporter to TLR assay is important to increase signal-to-noise ratio in the readout of a particular assay.

## Brief Description of the Figures

This application includes examples which refer to figures or other drawings. It is to be understood that the referenced figures are illustrative only and are not essential to the enablement of the claimed invention.

Figure 1 is two paired bar graphs showing (A) the induction of NF-κB and (B) the amount of IL-8 produced by 293 fibroblast cells transfected with human TLR9 in response to exposure to various stimuli, including CpG-ODN, GpC-ODN, LPS, and medium.

Figure 2 is a bar graph showing the induction of NF-κB produced by 293 fibroblast cells transfected with murine TLR9 in response to exposure to various

stimuli, including CpG-ODN, methylated CpG-ODN (Me-CpG-ODN), GpC-ODN, LPS, and medium.

Figure 3 is a series of gel images depicting the results of reverse transcriptase-polymerase chain reaction (RT-PCR) assays for murine TLR9 (mTLR9), human TLR9 (hTLR9), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in untransfected control 293 cells, 293 cells transfected with mTLR9 (293-mTLR9), and 293 cells transfected with hTLR9 (293-hTLR9).

Figure 4 is a graph showing the degree of induction of NF-κB-luc by various stimuli in stably transfected 293-hTLR9 cells.

Figure 5 is a graph showing the degree of induction of NF-κB-luc by various stimuli in stably transfected 293-mTLR9 cells.

Figure 6 is a graph showing fold induction of response as a function of concentration for a series of four related immunostimulatory nucleic acids contacted with human 293 fibroblast cells stably transfected with murine TLR9 and NF-κB-luc. Concentrations listed correspond to EC50 for each ligand.

Figure 7 is a graph showing kinetics of EC50 determinations for a series of five immunostimulatory nucleic acids contacted with human 293 fibroblast cells stably transfected with murine TLR9 and NF-κB-luc.

Figure 8 is a graph showing kinetics of EC50 determinations for the same series of five immunostimulatory nucleic acids as in Figure 7 contacted with human 293 fibroblast cells stably transfected with human TLR9 and NF-κB-luc.

Figure 9 is a graph showing kinetics of maximal activity (fold induction of response) for the same series of five immunostimulatory nucleic acids as in Figure 7 contacted with human 293 fibroblast cells stably transfected with murine TLR9 and NF-κB-luc.

Figure 10 is a graph showing kinetics of maximal activity (fold induction of response) for the same series of five immunostimulatory nucleic acids as in Figure 7 contacted with human 293 fibroblast cells stably transfected with human TLR9 and NF-κB-luc.

Figure 11 is a bar graph showing fold induction of response as measured using various luciferase reporter constructs (NF-κB-luc, IP-10-luc, RANTES-luc, ISRE-luc,

and IL-8-luc) in combination with TLR7, TLR8, and TLR9, each TLR contacted with a specific reference TLR ligand.

## Detailed Description of the Invention

The invention in certain aspects provides screening methods useful for the identification, characterization, and optimization of immunostimulatory compounds, including but not limited to immunostimulatory nucleic acids and immunostimulatory small molecules, as well as assays for the identification and optimization of agonists and antagonists of TLR3 signaling. The methods according to the invention include both cell-based and cell-free assays. In certain preferred embodiments the screening methods are performed in a high throughput manner. The methods can be used to screen libraries of compounds for their ability to modulate immune activation that involves TLR3 signaling.

In one aspect the invention provides a screening method for identifying an immunostimulatory compound. The method according to this aspect of the invention involves the steps of (a) contacting a functional TLR3 with a test compound under conditions which, in absence of the test compound, permit a negative control response mediated by a TLR3 signal transduction pathway; (b) detecting a test response mediated by the TLR3 signal transduction pathway; and (c) determining the test compound is an immunostimulatory compound when the test response exceeds the negative control response. In a second aspect the invention provides a screening method for identifying an immunostimulatory compound. The method according to this aspect of the invention involves the steps of (a) contacting a functional TLR3 with a test compound under conditions which, in presence of a reference immunostimulatory compound, permit a reference response mediated by a TLR3 signal transduction pathway; (b) detecting a test response mediated by the TLR3 signal transduction pathway; and (c) determining the test compound is an immunostimulatory compound when the test response equals or exceeds the reference response. It will be appreciated that these two aspects of the invention differ in that one involves comparison of the test compound against a negative control and the other involves comparison of the test compound against a positive control.

For these and other aspects of the invention, the TLR3 is preferably a mammalian TLR3, such as human TLR3 or mouse TLR3. Nucleotide and amino acid sequences for human TLR3 and murine TLR3 have previously been described. The nucleotide sequence for human TLR3 cDNA can be found as GenBank accession no. NM_003265 (SEQ ID NO:1), and the deduced amino acid sequence for human TLR3, encompassing 904 amino acids, can be found as GenBank accession nos NP_003256 (SEQ ID NO:2). The nucleotide sequence for murine TLR3 cDNA can be found as GenBank accession no. AF355152 (SEQ ID NO:3), and the deduced amino acid sequence for murine TLR3, encompassing 905 amino acids, can be found as GenBank accession no. AAK26117 (SEQ ID NO:4).

As used herein, a "functional TLR3" shall refer to a polypeptide, including a full length naturally occurring TLR3 polypeptide as described above, which specifically binds a TLR3 ligand and signals via a Toll/interleukin-1 receptor (TIR) domain. In addition to full length naturally occurring TLR3, a functional TLR3 thus also refers to allelic variants, fusion proteins, and truncated versions of the same, provided the polypeptide specifically binds a TLR3 ligand and signals via a TIR domain. In a preferred embodiment, the functional TLR3 includes a human TLR3 extracellular domain having an amino acid sequence provided by amino acids 38-707 according to SEQ ID NO:2. In another preferred embodiment, the functional TLR3 includes a murine TLR3 extracellular domain having an amino acid sequence provided by amino acids 39-708 according to SEQ ID NO:4. Preferably, the functional TLR3 signals through a TIR domain of TLR3.

In certain embodiments of this and other aspects of the invention, the functional TLR3 is expressed, either naturally or artifically, in a cell. In some embodiments, a cell expressing TLR3 for use in the methods of the invention expresses TLR3 and no other TLR. Alternatively, in some embodiments a cell expressing TLR3 for use in the methods of the invention expresses both TLR3 and at least one other TLR, e.g., TLR7, TLR8, or TLR9. In one embodiment the cell is an isolated mammalian cell that naturally expresses functional TLR3. Cells and tissues known to express TLR3 include dendritic cells (DCs), intraepithelial cells, and placenta. Muzio M et al. (2000) *J Immunol* 164:5998-6004; Cario E et al. (2000) *Infect Immun* 68:7010-7; Rock FL et al. (1998) *Proc Natl Acad Sci USA* 95:588-93. The term "isolated" as used herein, with

reference to a cell or to a compound, means substantially free of or separated from components with which the cell or compound is normally associated in nature, e.g., other cells, nucleic acids, proteins, lipids, carbohydrates or *in vivo* systems to an extent practical and appropriate for its intended use.

In another embodiment the cell can be one that, as it occurs in nature, is not capable of expressing TLR3 but which is rendered capable of expressing TLR3 through the artificial introduction of an expression vector for TLR3. Examples of cell lines lacking TLR3 include, but are not limited to, human 293 fibroblasts (ATCC CRL-1573) and HEp-2 human epithelial cells (ATCC CCL-23). Examples of cell lines lacking TLR9 include, but are not limited to, human 293 fibroblasts (ATCC CRL-1573), MonoMac-6, THP-1, U937, CHO, and any TLR9 knock-out. Typically the cell, whether it is capable of expressing TLR3 naturally or artificially, preferably has all the necessary elements for signal transduction initiated through the the TLR3 receptor. For example, it is believed that TLR9 signaling requires the adapter protein MyD88 in an early step of signal transduction. In contrast, TLR3 appears not to require MyD88 but may require other factors further downstream, e.g., factors that induce mitogen-activated protein kinase (MAPK) and factors downstream of MAPK.

When indicated, introduction of a particular TLR into a cell or cell line is preferably accomplished by transient or stable transfection of the cell or cell line with a TLR-encoding nucleic acid sequence operatively linked to a gene expression sequence (as described herein). For example, a cell artificially induced to express TLR3 for use in the methods of the invention includes a cell that has been transiently or stably transfected with a TLR3 expression vector. Any suitable method of transient or stable transfection can be employed for this purpose.

An expression vector for TLR3 will include at least a nucleotide sequence coding for a functional TLR3 polypeptide, operably linked to a gene expression sequence which can direct the expression of the TLR3 nucleic acid within a eukaryotic or prokaryotic cell. A "gene expression sequence" is any regulatory nucleotide sequence, such as a promoter sequence or promoter-enhancer combination, which facilitates the efficient transcription and translation of the nucleic acid to which it is operably linked. With respect to TLR3 nucleic acid, the "gene expression sequence" is any regulatory nucleotide sequence, such as a promoter sequence or promoter-enhancer

combination, which facilitates the efficient transcription and translation of the TLR3 nucleic acid to which it is operably linked. The gene expression sequence may, for example, be a mammalian or viral promoter, such as a constitutive or inducible promoter. Constitutive mammalian promoters include, but are not limited to, the promoters for the following genes: hypoxanthine phosphoribosyl transferase (HPRT), adenosine deaminase, pyruvate kinase, β-actin promoter, and other constitutive promoters. Exemplary viral promoters which function constitutively in eukaryotic cells include, for example, promoters from the simian virus (e.g., SV40), papillomavirus, adenovirus, human immunodeficiency virus (HIV), Rous sarcoma virus (RSV), cytomegalovirus (CMV), the long terminal repeats (LTR) of Moloney murine leukemia virus and other retroviruses, and the thymidine kinase (TK) promoter of herpes simplex virus. Other constitutive promoters are known to those of ordinary skill in the art. The promoters useful as gene expression sequences of the invention also include inducible promoters. Inducible promoters are expressed in the presence of an inducing agent. For example, the metallothionein (MT) promoter is induced to promote transcription and translation in the presence of certain metal ions. Other inducible promoters are known to those of ordinary skill in the art.

In general, the gene expression sequence shall include, as necessary, 5' non-transcribing and 5' non-translating sequences involved with the initiation of transcription and translation, respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribing sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined TLR3 nucleic acid. The gene expression sequences optionally include enhancer sequences or upstream activator sequences as desired.

Generally a nucleic acid coding sequence and a gene expression sequence are said to be "operably linked" when they are covalently linked in such a way as to place the transcription and/or translation of the nucleic acid coding sequence under the influence or control of the gene expression sequence. Thus the TLR3 nucleic acid sequence and the gene expression sequence are said to be "operably linked" when they are covalently linked in such a way as to place the transcription and/or translation of the TLR3 coding sequence under the influence or control of the gene expression sequence. If it is desired that the TLR3 sequence be translated into a functional protein, two DNA

sequences are said to be operably linked if induction of a promoter in the 5' gene expression sequence results in the transcription of the TLR3 sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the TLR3 sequence, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a gene expression sequence would be operably linked to a TLR3 nucleic acid sequence if the gene expression sequence were capable of effecting transcription of that TLR3 nucleic acid sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

In certain embodiments a TLR expression vector is constructed so as to permit tandem expression of two distinct TLRs, e.g., both TLR3 and a second TLR. Such a tandem expression vector can be used when it is desired to express two TLRs using a single transformation or transfection. Alternatively, a TLR3 expression vector can be used in conjunction with a second expression vector constructed so as to permit expression of a second TLR.

The screening assays can have any of a number of possible readout systems based upon a TLR/IL-1R signal transduction pathway. In preferred embodiments, the readout for the screening assay is based on the use of native genes or, alternatively, transfected or otherwise artificially introduced reporter gene constructs which are responsive to the TLR/IL-1R signal transduction pathway involving MyD88, TRAF, p38, and/or ERK. Häcker H et al. (1999) *EMBO J* 18:6973-82. These pathways activate kinases including κB kinase complex and c-Jun N-terminal kinases. Thus reporter genes and reporter gene constructs particularly useful for the assays include, e.g., a reporter gene operatively linked to a promoter sensitive to NF-κB. Examples of such promoters include, without limitation, those for NF-κB, IL-1β, IL-6, IL-8, IL-12 p40, CD80, CD86, and TNF-α. The reporter gene operatively linked to the TLR-sensitive promoter can include, without limitation, an enzyme (e.g., luciferase, alkaline phosphatase, β-galactosidase, chloramphenicol acetyltransferase (CAT), etc.), a bioluminescence marker (e.g., green-fluorescent protein (GFP, U.S. patent 5,491,084), etc.), a surface-expressed molecule (e.g., CD25), and a secreted molecule (e.g., IL-8, IL-12 p40, TNF-α). In certain preferred embodiments the reporter is selected from IL-

8, TNF-α, NF-κB-luciferase (NF-κB-luc; Häcker H et al. (1999) *EMBO J* 18:6973-82), IL-12 p40-luc (Murphy TL et al. (1995) *Mol Cell Biol* 15:5258-67), and TNF-luc (Häcker H et al. (1999) *EMBO J* 18:6973-82). In assays relying on enzyme activity readout, substrate can be supplied as part of the assay, and detection can involve measurement of chemiluminescence, fluorescence, color development, incorporation of radioactive label, drug resistance, or other marker of enzyme activity. For assays relying on surface expression of a molecule, detection can be accomplished using flow cytometry (FACS) analysis or functional assays. Secreted molecules can be assayed using enzyme-linked immunosorbent assay (ELISA) or bioassays. These and other suitable readout systems are well known in the art and are commercially available.

Thus a cell expressing a functional TLR3 and useful for the methods of the invention has, in some embodiments, an expression vector comprising an isolated nucleic acid which encodes a reporter construct useful for detecting TLR signaling. The expression vector comprising an isolated nucleic acid which encodes a reporter construct useful for detecting TLR signaling can include a reporter gene under control of a minimal promoter responsive to a transcription factor believed by the applicant to be activated as a consequence of TLR3 signaling. Examples of such minimal promoters include, without limitation, promoters for the following genes: AP1, NF-κB, ATF2, IRF3, and IRF7. In other embodiments the expression vector comprising an isolated nucleic acid which encodes a reporter construct useful for detecting TLR signaling can include a gene under control of a promoter response element selected from IL-6, IL-8, IL-12 p40 subunit, a type 1 IFN, RANTES, TNF, IP-10, I-TAC, and ISRE. The promoter response element generally will be present in multiple copies, e.g., as tandem repeats. For example, an ISRE-luciferase reporter construct useful in the invention is available from Stratagene (catalog no. 219092) and includes a 5x ISRE tandem repeat joined to a TATA box upstream of a luciferase reporter gene. As discussed further elsewhere herein, the reporter itself can be any gene product suitable for detection by methods recognized in the art. Such methods for detection can include, for example, measurement of spontaneous or stimulated light emission, enzyme activity, expression of a soluble molecule, expression of a cell surface molecule, etc.

As mentioned above, the functional TLR3 is contacted with a test compound in order to identify an immunostimulatory compound. An immunostimulatory compound

is a natural or synthetic compound that is capable of inducing an immune response when contacted with an immune cell. In the context of the methods of the invention, an immunostimulatory compound refers to a natural or synthetic compound that is capable of inducing an immune response when contacted with an immune cell expressing a functional TLR3 polypeptide. Preferably the immune response is or involves activation of a TLR3 signal transduction pathway. Thus immunostimulatory compounds identified and characterized using the methods of the invention specifically include TLR3 ligands, i.e., compounds which selectively bind to TLR3 and induce a TLR3 signal transduction pathway. Immunostimulatory compounds in general include but are not limited to nucleic acids, including oligonucleotides and polynucleotides; oligopeptides; polypeptides; lipids, including lipopolysaccharides; carbohydrates, including oligosaccharides and polysaccharides; and small molecules. Accordingly, a "test compound" refers to nucleic acids, including oligonucleotides and polynucleotides; oligopeptides; polypeptides; lipids, including lipopolysaccharides; carbohydrates, including oligosaccharides and polysaccharides; and small molecules. Test compounds include compounds with known biological activity as well as compounds without known biological activity.

A "reference immunostimulatory compound" refers to an immunostimulatory compound that characteristically induces an immune response when contacted with an immune cell expressing a functional TLR polypeptide. In the screening methods of the invention, the reference immunositmulatory compound is a natural or synthetic compound that that characteristically induces an immune response when contacted with an immune cell expressing a functional TLR3 polypeptide. Preferably the immune response is or involves activation of a TLR3 signal transduction pathway. Thus a reference immunostimulatory compound will characteristically induce a reference response mediated by a TLR3 signal transduction pathway when contacted with a functional TLR3 under suitable conditions. The reference response can be measured according to any of the methods described herein. Importantly, a reference immunostimulatory compound specifically includes a test compound identified as an immunostimulatory compound according to any one of the methods of the invention. Therefore a reference immunostimulatory compound can be a nucleic acid, including oligonucleotides and polynucleotides; an oligopeptide; a polypeptide; a lipid, including

lipopolysaccharides; a carbohydrate, including oligosaccharides and polysaccharides; or a small molecule.

Small molecules include naturally occurring, synthetic, and semisynthetic organic and organometallic compounds with molecular weight less than about 1.5 kDa. Examples of small molecules include most drugs, subunits of polymeric materials, and analogs and derivatives thereof.

A "nucleic acid" as used herein with respect to test compounds and reference compounds used in the methods of the invention, shall refer to any polymer of two or more individual nucleoside or nucleotide units. Typically individual nucleoside or nucleotide units will include any one or combination of deoxyribonucleosides, ribonucleosides, deoxyribonucleotides, and ribonucleotides. The individual nucleotide or nucleoside units of the nucleic acid can be naturally occurring or not naturally occurring. For example, the individual nucleotide units can include deoxyadenosine, deoxycytidine, deoxyguanosine, thymidine, and uracil. In addition to naturally occurring 2'-deoxy and 2'-hydroxyl forms, individual nucleosides also include synthetic nucleosides having modified base moieties and/or modified sugar moieties, e.g., as described in Uhlmann E et al. (1990) *Chem Rev* 90:543-84. The linkages between individual nucleotide or nucleoside units can be naturally occurring or not naturally occurring. For example, the linkages can be phosphodiester, phosphorothioate, phosphorodithioate, phosphoramidate, as well as peptide linkages and other covalent linkages, known in the art, suitable for joining adjacent nucleoside or nucleotide units. The nucleic acid test compounds and nucleic acid reference compounds typically range in size from 3-4 units to a few tens of units, e.g., 18-40 units.

The substituted purines and pyrimidines of the ISNAs include standard purines and pyrimidines such as cytosine as well as base analogs such as C-5 propyne substituted bases. Wagner RW et al. (1996) *Nat Biotechnol* 14:840-4. Purines and pyrimidines include but are not limited to adenine, cytosine, guanine, thymine, 5-methylcytosine, 2-aminopurine, 2-amino-6-chloropurine, 2,6-diaminopurine, hypoxanthine, and other naturally and non-naturally occurring nucleobases, substituted and unsubstituted aromatic moieties.

Libraries of compounds that can be used as test compounds are available from various commercial suppliers, and they can be made to order using techniques well

known in the art, including combinatorial chemistry techniques. Especially in combination with high throughput screening methods, such methods including in particular automated multichannel methods of screening, large libraries of test compounds can be screened according to the methods of the invention. Large libraries can include hundreds, thousands, tens of thousands, hundreds of thousands, and even millions of compounds.

Thus in preferred embodiments, the methods for screening test compounds can be performed on a large scale and with high throughput by incorporating, e.g., an array-based assay system and at least one automated or semi-automated step. For example, the assays can be set up using multiple-well plates in which cells are dispensed in individual wells and reagents are added in a systematic manner using a multiwell delivery device suited to the geometry of the multiwell plate. Manual and robotic multiwell delivery devices suitable for use in a high throughput screening assay are well known by those skilled in the art. Each well or array element can be mapped in a one-to-one manner to a particular test condition, such as the test compound. Readouts can also be performed in this multiwell array, preferably using a multiwell plate reader device or the like. Examples of such devices are well known in the art and are available through commercial sources. Sample and reagent handling can be automated to further enhance the throughput capacity of the screening assay, such that dozens, hundreds, thousands, or even millions of parallel assays can be performed in a day or in a week. Fully robotic systems are known in the art for applications such as generation and analysis of combinatorial libraries of synthetic compounds. See, for example, U.S. patents 5,443,791 and 5,708,158.

A "CpG nucleic acid" or a "CpG immunostimulatory nucleic acid" as used herein is a nucleic acid containing at least one unmethylated CpG dinucleotide (cytosine-guanine dinucleotide sequence, i.e. "CpG DNA" or DNA containing a 5' cytosine followed by 3' guanine and linked by a phosphate bond) and activates a component of the immune system. The entire CpG nucleic acid can be unmethylated or portions may be unmethylated but at least the C of the 5' CG 3' must be unmethylated.

In one embodiment a CpG nucleic acid is represented by at least the formula:

$$5'\text{-}N_1X_1CGX_2N_2\text{-}3'$$

wherein $X_1$ and $X_2$ are nucleotides, N is any nucleotide, and $N_1$ and $N_2$ are nucleic acid sequences composed of from about 0-25 N's each. In some embodiments $X_1$ is adenine, guanine, or thymine and/or $X_2$ is cytosine, adenine, or thymine. In other embodiments $X_1$ is cytosine and/or $X_2$ is guanine.

Examples of CpG nucleic acids according to the invention include but are not limited to those listed in Table 1.

Table 1. Exemplary CpG Nucleic Acids

| | |
|---|---|
| AACGTTCT | |
| AAGCGAAAATGAAATTGACT | SEQ ID NO:39 |
| ACCATGGACGAACTGTTTCCCCTC | SEQ ID NO:40 |
| ACCATGGACGACCTGTTTCCCCTC | SEQ ID NO:41 |
| ACCATGGACGAGCTGTTTCCCCTC | SEQ ID NO:42 |
| ACCATGGACGATCTGTTTCCCCTC | SEQ ID NO:43 |
| ACCATGGACGGTCTGTTTCCCCTC | SEQ ID NO:44 |
| ACCATGGACGTACTGTTTCCCCTC | SEQ ID NO:45 |
| ACCATGGACGTTCTGTTTCCCCTC | SEQ ID NO:46 |
| AGCGGGGGCGAGCGGGGGCG | SEQ ID NO:47 |
| AGCTATGACGTTCCAAGG | SEQ ID NO:48 |
| ATCGACTCTCGAGCGTTCTC | SEQ ID NO:49 |
| ATGACGTTCCTGACGTT | SEQ ID NO:50 |
| ATGGAAGGTCCAACGTTCTC | SEQ ID NO:51 |
| ATGGAAGGTCCAGCGTTCTC | SEQ ID NO:52 |
| ATGGACTCTCCAGCGTTCTC | SEQ ID NO:53 |
| ATGGAGGCTCCATCGTTCTC CAACGTT | SEQ ID NO:54 |
| CACGTTGAGGGGCAT | SEQ ID NO:55 |
| CAGGCATAACGGTTCCGTAG CCAACGTT | SEQ ID NO:56 |
| CTGATTTCCCCGAAATGATG | SEQ ID NO:57 |
| GAGAACGATGGACCTTCCAT | SEQ ID NO:58 |
| GAGAACGCTCCAGCACTGAT | SEQ ID NO:59 |
| GAGAACGCTCGACCTTCCAT | SEQ ID NO:60 |
| GAGAACGCTCGACCTTCGAT | SEQ ID NO:61 |
| GAGAACGCTGGACCTTCCAT | SEQ ID NO:62 |
| GATTGCCTGACGTCAGAGAG | SEQ ID NO:63 |
| GCATGACGTTGAGCT | SEQ ID NO:64 |
| GCGGCGGGCGGCGCGCGCCC | SEQ ID NO:65 |
| GCGTGCGTTGTCGTTGTCGTT | SEQ ID NO:66 |
| GCTAGACGTTAGCGT | SEQ ID NO:67 |
| GCTAGACGTTAGTGT | SEQ ID NO:68 |
| GCTAGATGTTAGCGT | SEQ ID NO:69 |
| GCTTGATGACTCAGCCGGAA | SEQ ID NO:70 |
| GGAATGACGTTCCCTGTG | SEQ ID NO:71 |

| | |
|---|---|
| GGGGTCAACGTTGACGGGG | SEQ ID NO:72 |
| GGGGTCAGTCTTGACGGGG | SEQ ID NO:73 |
| GTCCATTTCCCGTAAATCTT | SEQ ID NO:74 |
| GTCGCT | |
| GTCGTT | |
| TACCGCGTGCGACCCTCT | SEQ ID NO:75 |
| TCAACGTC | |
| TCAACGTT | |
| TCAGCGCT | |
| TCAGCGTGCGCC | SEQ ID NO:76 |
| TCATCGAT | |
| TCCACGACGTTTTCGACGTT | SEQ ID NO:77 |
| TCCATAACGTTCCTGATGCT | SEQ ID NO:78 |
| TCCATAGCGTTCCTAGCGTT | SEQ ID NO:79 |
| TCCATCACGTGCCTGATGCT | SEQ ID NO:80 |
| TCCATGACGGTCCTGATGCT | SEQ ID NO:81 |
| TCCATGACGTCCCTGATGCT | SEQ ID NO:82 |
| TCCATGACGTGCCTGATGCT | SEQ ID NO:83 |
| TCCATGACGTTCCTGACGTT | SEQ ID NO:84 |
| TCCATGACGTTCCTGATGCT | SEQ ID NO:18 |
| TCCATGCCGGTCCTGATGCT | SEQ ID NO:85 |
| TCCATGCGTGCGTGCGTTTT | SEQ ID NO:86 |
| TCCATGCGTTGCGTTGCGTT | SEQ ID NO:87 |
| TCCATGGCGGTCCTGATGCT | SEQ ID NO:88 |
| TCCATGTCGATCCTGATGCT | SEQ ID NO:89 |
| TCCATGTCGCTCCTGATGCT | SEQ ID NO:90 |
| TCCATGTCGGTCCTGATGCT | SEQ ID NO:91 |
| TCCATGTCGGTCCTGCTGAT | SEQ ID NO:92 |
| TCCATGTCGTCCCTGATGCT | SEQ ID NO:93 |
| TCCATGTCGTTCCTGATGCT | SEQ ID NO:94 |
| TCCATGTCGTTCCTGTCGTT | SEQ ID NO:95 |
| TCCATGTCGTTTTTGTCGTT | SEQ ID NO:96 |
| TCCTGACGTTCCTGACGTT | SEQ ID NO:97 |
| TCCTGTCGTTCCTGTCGTT | SEQ ID NO:98 |
| TCCTGTCGTTCCTTGTCGTT | SEQ ID NO:99 |
| TCCTGTCGTTTTTTGTCGTT | SEQ ID NO:100 |
| TCCTTGTCGTTCCTGTCGTT | SEQ ID NO:101 |
| TCGATCGGGGCGGGGCGAGC | SEQ ID NO:102 |
| TCGTCGCTGTCTCCGCTTCTT | SEQ ID NO:103 |
| TCGTCGCTGTCTCCGCTTCTTCTTGCC | SEQ ID NO:104 |
| TCGTCGCTGTCTGCCCTTCTT | SEQ ID NO:105 |
| TCGTCGCTGTTGTCGTTTCTT | SEQ ID NO:106 |
| TCGTCGTCGTCGTT | SEQ ID NO:107 |
| TCGTCGTTGTCGTTGTCGTT | SEQ ID NO:108 |
| TCGTCGTTGTCGTTTTGTCGTT | SEQ ID NO:109 |
| TCGTCGTTTTGTCGTTTTGTCGTT | SEQ ID NO:15 |
| TCTCCCAGCGCGCGCCAT | SEQ ID NO:110 |
| TCTCCCAGCGGGCGCAT | SEQ ID NO:111 |

| | |
|---|---|
| TCTCCCAGCGTGCGCCAT TCTTCGAA | SEQ ID NO:112 |
| TGCAGATTGCGCAATCTGCA TGTCGCT TGTCGTT | SEQ ID NO:113 |
| TGTCGTTGTCGTT | SEQ ID NO:114 |
| TGTCGTTGTCGTTGTCGTT | SEQ ID NO:115 |
| TGTCGTTGTCGTTGTCGTTGTCGTT | SEQ ID NO:116 |
| TGTCGTTTGTCGTTTGTCGTT | SEQ ID NO:117 |

As used herein the term "response mediated by a TLR signal transduction pathway" refers to a response which is characteristic of an interaction between a TLR and an immunostimulatory compound that induces signaling events through the TLR. Such responses typically involve usual elements of Toll/IL-1R signaling, e.g., MyD88, TRAF, and IRAK molecules, although in the case of TLR3 the role of MyD88 is less clear than for other TLR family members. As demonstrated herein such responses include the induction of a gene under control of a specific promoter such as a NF-κB promoter, increases in particular cytokine levels, increases in particular chemokine levels etc. The gene under the control of the NF-κB promoter may be a gene which naturally includes an NF-κB promoter or it may be a gene in a construct in which an NF-κB promoter has been inserted. Genes which naturally include the NF-κB promoter include but are not limited to IL-8, IL-12 p40, NF-κB-luc, IL-12 p40-luc, and TNF-luc. Increases in cytokine levels may result from increased production or increased stability or increased secretion of the cytokines in response to the TLR-immunostimulatory compound interaction. Th1 cytokines include but are not limited to IL-2, IFN-γ, and IL-12. It has unexpectedly been discovered, according to the instant invention, that the promoter response element ISRE is directly activated as a result of signaling through the TLR3 signal transduction pathway, i.e., independent of IFN-γ production. Th2 cytokines include but are not limited to IL-4, IL-5, and IL-10. Chemokines of particular significance in the invention include but are not limited to CCL5 (RANTES), CXCL9 (Mig), CXCL10 (IP-10), and CXCL11 (I-TAC).

In another aspect the invention provides a screening method for identifying a compound that modulates TLR3 signaling activity. The method according to this aspect of the invention involves the steps of (a) contacting a functional TLR3 with a test compound and a reference immunostimulatory compound under conditions which,

in presence of the reference immunostimulatory compound alone, permit a reference response mediated by a TLR3 signal transduction pathway; (b) detecting a test-reference response mediated by the TLR3 signal transduction pathway; (c) determining the test compound is an agonist of TLR3 signaling activity when the test-reference response exceeds the reference response; and (d) determining the test compound is an antagonist of TLR3 signaling activity when the reference response exceeds the test-reference response. A test-reference response refers to a type of test response as determined when a test compound and a reference immunostimulatory compound are simultaneously contacted with the TLR3. When a test compound is neither an agonist nor an antagonist of TLR3 signaling activity, the test-reference response and the reference response are indistinguishable.

An agonist as used herein is a compound which causes an enhanced response of a TLR to a reference stimulus. The enhanced response can be additive or synergistic with respect to the response to the reference stimulus by itself. Furthermore, an agonist can work directly or indirectly to cause the enhanced response. Thus an agonist of TLR3 signaling activity as used herein is a compound which causes an enhanced response of a TLR to a reference stimulus.

An antagonist as used herein is a compound which causes a diminished response of a TLR to a reference stimulus. Furthermore, an antagonist can work directly or indirectly to cause the diminished response. Thus an antagonist of TLR3 signaling activity as used herein is a compound which causes a diminished response of a TLR to a reference stimulus.

In addition to identification and characterization of immunostimulatory compounds, agonists of TLR3 signaling, and antagonists of TLR3 signaling, the methods of the invention also permit optimization of lead compounds. Optimization of a lead compound involves an iterative application of a screening method of the invention, further including the steps of selecting the best candidate at any given stage or round in the screening and then substituting it as a benchmark or reference in a subsequent round of screening. This latter process can further include selection of parameters to modify in choosing and generating candidate test compounds to screen. For example, a lead compound from a particular round of screening can be used as a

basis to develop a focused library of new test compounds for use in a subsequent round of screening.

In another aspect the invention provides a screening method for identifying species specificity of an immunostimulatory compound. The method according to this aspect of the invention involves the steps of (a) measuring a first species-specific response mediated by a TLR3 signal transduction pathway when a functional TLR3 of a first species is contacted with a test compound; (b) measuring a second species-specific response mediated by the TLR3 signal transduction pathway when a functional TLR3 of a second species is contacted with the test compound; and (c) comparing the first species-specific response with the second species-specific response.

A species-specific TLR, including TLR3, is not limited to a human TLR, but rather can include a TLR derived from human or non-human sources. Examples of non-human sources include, but are not limited to, murine, rat, bovine, canine, feline, ovine, porcine, and equine. Other species include chicken and fish, e.g., aquaculture species.

The species-specific TLR, including TLR3, also is not limited to native TLR polypeptides. In certain embodiments the TLR can be, e.g., a chimeric TLR in which the extracellular domain and the cytoplasmic domain are derived from TLR polypeptides from different species. Such chimeric TLR polypeptides, as described above, can include, for example, a human TLR extracellular domain and a murine TLR cytoplasmic domain, each domain derived from the corresponding TLR of each species. In alternative embodiments, such chimeric TLR polypeptides can include chimeras created with different TLR splice variants or allotypes. Other chimeric TLR polypeptides useful for the screening methods of the invention include chimeric polypeptides created with a TLR of a first type, e.g., TLR3, and another TLR, e.g., TLR7, TLR8, or TLR9, of the same or another species as the TLR of the first type. Also contemplated are chimeric polypeptides which incorporate sequences derived from more than two polypeptides, e.g., an extracellular domain, a transmembrane domain, and a cytoplasmic domain all derived from different polypeptide sources, provided at least one such domain derives from a TLR3 polypeptide. As a further example, also contemplated are constructs such as include an extracellular domain of one TLR3, an intracellular domain of another TLR3, and a non-TLR reporter such as

luciferase, GFP, etc. Those of skill in the art will recognize how to design and generate DNA sequences coding for such chimeric TLR polypeptides.

It has also been discovered, according to the instant invention, that TLR-based screening assays, including but not limited to the TLR3-based assays described herein, are sensitive to parameters such as concentration of test compound, stability of test compound, kinetics of detection, and selection of reporter. These parameters can be optimized in order to derive the most information from a given screening assay. Importantly, the kinetics of detection appear to afford separation of types of information such as affinity of interaction and stability or duration of interaction. For example, measurements taken at earlier timepoints, e.g., after 6-8 hours of contact between TLR and test and/or reference compound, appear to reflect more information about affinity of interaction than do measurements obtained at later timepoints, e.g., after 16-24 or more hours of contact. In addition, while NF-κB-driven reporters are generally useful in TLR-based screening assays like those of the instant invention, in some instances a reporter other than an NF-κB-driven reporter will afford greater sensitivity. For example, the IL-8-luc reporter is significantly more sensitive to TLR7 and TLR8 than NF-κB-luc. Selection of reporter thus appears to be TLR-dependent, while parameters relating to kinetics and concentration appear to be more compound-dependent. Thus in performing the screening methods of the instant invention, it is expected that the methods will be enhance by inclusion of measurements obtained using at least two concentrations and two time points for each test compound. Typically at least three concentrations will be employed, spanning a two to three log-fold range of concentrations. Finer ranges of concentration can of course be employed under suitable circumstances, for instance based on results of an earlier screening performed using a wider initial range of concentrations.

The invention will be more fully understood by reference to the following examples. These examples, however, are merely intended to illustrate certain embodiments of the invention and are not to be construed to limit the scope of the invention.

## Examples

**Example 1. Expression Vectors for Human TLR3 (hTLR3) and Murine TLR3 (mTLR3)**

To create an expression vector for human TLR3, human TLR3 cDNA was amplified by the polymerase chain method (PCR) from a cDNA made from human 293 cells using the primers 5'-GAAACTCGAGCCACCATGAGACAGACTTTGCCTTGTATCTAC-3' (sense, SEQ ID NO:9) and 5'-GAAAGAATTCTTAATGTACAGAGTTTTTGGATCCAAG-3' (antisense, SEQ ID NO:10). The primers introduce Xho I and EcoRI restriction endonuclease sites at their 5' ends for use in subsequent cloning into the expression vector. The resulting amplication product fragment was cloned into pGEM-T Easy vector (Promega), isolated, cut with Xho I and EcoRI restriction endonucleases, ligated into an Xho I/EcoRI-digested pcDNA3.1 expression vector (Invitrogen). The insert was fully sequenced and translated into protein. The cDNA sequence corresponds to the published cDNA sequence for hTLR3, available as GenBank accession no. NM_003265 (SEQ ID NO:1). The open reading frame codes for a protein 904 amino acids long, having the sequence corresponding to GenBank accession no. NP_003256 (SEQ ID NO:2).

Table 2. cDNA Sequence for Human TLR3

(GenBank Accession No. NM_003265; SEQ ID NO:1)

```
gcggccgcgt cgacgaaatg tctggatttg gactaaagaa aaaaggaaag gctagcagtc   60
atccaacaga atcatgagac agactttgcc ttgtatctac ttttgggggg gccttttgcc  120
ctttgggatg ctgtgtgcat cctccaccac caagtgcact gttagccatg aagttgctga  180
ctgcagccac ctgaagttga ctcaggtacc cgatgatcta cccacaaaca taacagtgtt  240
gaaccttacc cataatcaac tcagaagatt accagccgcc aacttcacaa ggtatagcca  300
gctaactagc ttggatgtag gatttaacac catctcaaaa ctggagccag aattgtgcca  360
gaaacttccc atgttaaaag ttttgaacct ccagcacaat gagctatctc aactttctga  420
taaaaccttt gccttctgca cgaatttgac tgaactccat ctcatgtcca actcaatcca  480
gaaaattaaa aataatccct ttgtcaagca gaagaattta atcacattag atctgtctca  540
taatggcttg tcatctacaa aattaggaac tcaggttcag ctggaaaatc tccaagagct  600
tctattatca aacaataaaa ttcaagcgct aaaaagtgaa gaactggata tctttgccaa  660
ttcatcttta aaaaaattag agttgtcatc gaatcaaatt aaagagtttt ctccagggtg  720
ttttcacgca attggaagat tatttggcct ctttctgaac aatgtccagc tgggtcccag  780
ccttacagag aagctatgtt tggaattagc aaacacaagc attcggaatc tgtctctgag  840
taacagccag ctgtccacca ccagcaatac aactttcttg ggactaaagt ggacaaatct  900
cactatgctc gatctttcct acaacaactt aaatgtggtt ggtaacgatt cctttgcttg  960
gcttccacaa ctagaatatt tcttcctaga gtataataat atacagcatt tgttttctca 1020
ctctttgcac gggcttttca atgtgaggta cctgaatttg aaacggtctt ttactaaaca 1080
aagtatttcc cttgcctcac tccccaagat tgatgatttt tcttttcagt ggctaaaatg 1140
tttggagcac cttaacatgg aagataatga tattccaggc ataaaaagca atatgttcac 1200
aggattgata aacctgaaat acttaagtct atccaactcc tttacaagtt tgcgaacttt 1260
gacaaatgaa acatttgtat cacttgctca ttctccctta cacatactca acctaaccaa 1320
gaataaaatc tcaaaaatag agagtgatgc tttctcttgg ttgggccacc tagaagtact 1380
```

```
tgacctgggc cttaatgaaa ttgggcaaga actcacaggc caggaatgga gaggtctaga 1440
aaatattttc gaaatctatc tttcctacaa caagtacctg cagctgacta ggaactcctt 1500
tgccttggtc ccaagccttc aacgactgat gctccgaagg gtggccctta aaaatgtgga 1560
tagctctcct tcaccattcc agcctcttcg taacttgacc attctggatc taagcaacaa 1620
caacatagcc aacataaatg atgacatgtt ggagggtctt gagaaactag aaattctcga 1680
tttgcagcat aacaacttag cacggctctg gaaacacgca aaccctggtg gtcccattta 1740
tttcctaaag ggtctgtctc acctccacat ccttaacttg gagtccaacg gctttgacga 1800
gatcccagtt gaggtcttca aggatttatt tgaactaaag atcatcgatt taggattgaa 1860
taatttaaac acacttccag catctgtctt taataatcag gtgtctctaa agtcattgaa 1920
ccttcagaag aatctcataa catccgttga gaagaaggtt ttcgggccag ctttcaggaa 1980
cctgactgag ttagatatgc gctttaatcc ctttgattgc acgtgtgaaa gtattgcctg 2040
gtttgttaat tggattaacg agacccatac caacatccct gagctgtcaa gccactacct 2100
ttgcaacact ccacctcact atcatgggtt cccagtgaga ctttttgata catcatcttg 2160
caaagacagt gccccctttg aactcttttt catgatcaat accagtatcc tgttgatttt 2220
tatctttatt gtacttctca tccactttga gggctggagg atatcttttt attggaatgt 2280
ttcagtacat cgagttcttg gtttcaaaga aatagacaga cagacagaac agtttgaata 2340
tgcagcatat ataattcatg cctataaaga taaggattgg gtctgggaac atttctcttc 2400
aatggaaaag gaagaccaat ctctcaaatt ttgtctggaa gaaagggact ttgaggcggg 2460
tgtttttgaa ctagaagcaa ttgttaacag catcaaaaga agcagaaaaa ttatttttgt 2520
tataacacac catctattaa aagacccatt atgcaaaaga ttcaaggtac atcatgcagt 2580
tcaacaagct attgaacaaa atctggattc cattatattg gttttccttg aggagattcc 2640
agattataaa ctgaaccatg cactctgttt gcgaagagga atgtttaaat ctcactgcat 2700
cttgaactgg ccagttcaga aagaacggat aggtgccttt cgtcataaat tgcaagtagc 2760
acttggatcc aaaaactctg tacattaaat ttatttaaat attcaattag caaaggagaa 2820
actttctcaa tttaaaaagt tctatggcaa atttaagttt tccataaagg tgttataatt 2880
tgtttattca tatttgtaaa tgattatatt ctatcacaat tacatctctt ctaggaaaat 2940
gtgtctcctt atttcaggcc tatttttgac aattgactta attttaccca aaataaaaca 3000
tataagcacg caaaaaaaaa aaaaaaaaa                                    3029
```

Table 3. Amino Acid Sequence for Human TLR3

(GenBank Accession No. NP_003256; SEQ ID NO:2)

```
MRQTLPCIYF WGGLLPFGML CASSTTKCTV SHEVADCSHL KLTQVPDDLP TNITVLNLTH   60
NQLRRLPAAN FTRYSQLTSL DVGFNTISKL EPELCQKLPM LKVLNLQHNE LSQLSDKTFA  120
FCTNLTELHL MSNSIQKIKN NPFVKQKNLI TLDLSHNGLS STKLGTQVQL ENLQELLLSN  180
NKIQALKSEE LDIFANSSLK KLELSSNQIK EFSPGCFHAI GRLFGLFLNN VQLGPSLTEK  240
LCLELANTSI RNLSLSNSQL STTSNTTFLG LKWTNLTMLD LSYNNLNVVG NDSFAWLPQL  300
EYFFLEYNNI QHLFSHSLHG LFNVRYLNLK RSFTKQSISL ASLPKIDDFS FQWLKCLEHL  360
NMEDNDIPGI KSNMFTGLIN LKYLSLSNSF TSLRTLTNET FVSLAHSPLH ILNLTKNKIS  420
KIESDAFSWL GHLEVLDLGL NEIGQELTGQ EWRGLENIFE IYLSYNKYLQ LTRNSFALVP  480
SLQRLMLRRV ALKNVDSSPS PFQPLRNLTI LDLSNNNIAN INDDMLEGLE KLEILDLQHN  540
NLARLWKHAN PGGPIYFLKG LSHLHILNLE SNGFDEIPVE VFKDLFELKI IDLGLNNLNT  600
LPASVFNNQV SLKSLNLQKN LITSVEKKVF GPAFRNLTEL DMRFNPFDCT CESIAWFVNW  660
INETHTNIPE LSSHYLCNTP PHYHGFPVRL FDTSSCKDSA PFELFFMINT SILLIFIFIV  720
LLIHFEGWRI SFYWNVSVHR VLGFKEIDRQ TEQFEYAAYI IHAYKDKDWV WEHFSSMEKE  780
DQSLKFCLEE RDFEAGVFEL EAIVNSIKRS RKIIFVITHH LLKDPLCKRF KVHHAVQQAI  840
EQNLDSIILV FLEEIPDYKL NHALCLRRGM FKSHCILNWP VQKERIGAFR HKLQVALGSK  900
NSVH                                                               904
```

Corresponding nucleotide and amino acid sequences for murine TLR3 (mTLR3) are known. The nucleotide sequence of mTLR3 cDNA has been reported as GenBank accession no. AF355152, and the amino acid sequence of mTLR3 has been reported as GenBank accession no. AAK26117.

Table 4. cDNA Sequence for Murine TLR3

(GenBank Accession No. AF355152; SEQ ID NO:3)

```
tagaatatga tacagggatt gcacccataa tctgggctga atcatgaaag ggtgttcctc   60
ttatctaatg tactcctttg ggggactttt gtccctatgg attcttctgg tgtcttccac  120
aaaccaatgc actgtgagat acaacgtagc tgactgcagc catttgaagc taacacacat  180
acctgatgat cttccctcta acataacagt gttgaatctt actcacaacc aactcagaag  240
attaccacct accaacttta caagatacag ccaacttgct atcttggatg caggatttaa  300
ctccatttca aaactggagc cagaactgtg ccaaatactc cctttgttga aagtattgaa  360
cctgcaacat aatgagctct ctcagatttc tgatcaaacc tttgtcttct gcacgaacct  420
gacagaactc gatctaatgt ctaactcaat acacaaaatt aaaagcaacc ctttcaaaaa  480
ccagaagaat ctaatcaaat tagatttgtc tcataatggt ttatcatcta caaagttggg  540
aacgggggtc caactggaga acctccaaga actgctctta gcaaaaaata aaatccttgc  600
gttgcgaagt gaagaacttg agtttcttgg caattcttct ttacgaaagt tggacttgtc  660
atcaaatcca cttaaagagt tctccccggg gtgtttccag acaattggca agttattcgc  720
cctcctcttg aacaacgccc aactgaaccc ccacctcaca gagaagcttt gctgggaact  780
ttcaaacaca agcatccaga atctctctct ggctaacaac cagctgctgg ccaccagcga  840
gagcactttc tctgggctga agtggacaaa tctcacccag ctcgatcttt cctacaacaa  900
cctccatgat gtcggcaacg gttccttctc ctatctccca agcctgaggt atctgtctct  960
ggagtacaac aatatacagc gtctgtcccc tcgctctttt tatggactct ccaacctgag 1020
gtacctgagt ttgaagcgag catttactaa gcaaagtgtt tcacttgctt cacatcccaa 1080
cattgacgat ttttcctttc aatggttaaa atatttggaa tatctcaaca tggatgacaa 1140
taatattcca agtaccaaaa gcaatacctt cacgggattg gtgagtctga agtacctaag 1200
tctttccaaa actttcacaa gtttgcaaac tttaacaaat gaaacatttg tgtcacttgc 1260
tcattctccc ttgctcactc tcaacttaac gaaaaatcac atctcaaaaa tagcaaatgg 1320
tactttctct tggttaggcc aactcaggat acttgatctc ggccttaatg aaattgaaca 1380
aaaactcagc ggccaggaat ggagaggtct gagaaatata tttgagatct acctatccta 1440
taacaaatac ctccaactgt ctaccagttc ctttgcattg gtccccagcc ttcaaagact 1500
gatgctcagg agggtggccc ttaaaaatgt ggatatctcc ccttcacctt tccgccctct 1560
tcgtaacttg accattctgg acttaagcaa caacaacata gccaacataa atgaggactt 1620
gctggagggt cttgagaatc tagaaatcct ggattttcag cacaataact tagccaggct 1680
ctggaaacgc gcaaaccccg gtggtcccgt taatttcctg aaggggctgt ctcacctcca 1740
catcttgaat ttagagtcca acggcttaga tgaaatccca gtcggggttt tcaagaactt 1800
attcgaacta aagagcatca atctaggact gaataactta aacaaacttg aaccattcat 1860
ttttgatgac cagacatctc taaggtcact gaacctccag aagaacctca taacatctgt 1920
tgagaaggat gttttcgggc cgccttttca aaacctgaac agtttagata tgcgcttcaa 1980
tccgttcgac tgcacgtgtg aaagtatttc ctggtttgtt aactggatca accagaccca 2040
cactaatatc tttgagctgt ccactcacta cctctgtaac actccacatc attattatgg 2100
cttccccctg aagcttttcg atacatcatc ctgtaaagac agcgcccccc ttgaactcct 2160
cttcataatc agcaccagta tgctcctggt ttttatactt gtggtactgc tcattcacat 2220
cgagggctgg aggatctctt tttactggaa tgtttcagtg catcggattc ttggtttcaa 2280
ggaaatagac acacaggctg agcagtttga atatacagcc tacataattc atgcccataa 2340
agacagagac tgggtctggg aacatttctc cccaatggaa gaacaagacc aatctctcaa 2400
attttgccta gaagaaaggg actttgaagc aggcgtcctt ggacttgaag caattgttaa 2460
tagcatcaaa agaagccgaa aaatcatttt cgttatcaca caccatttat taaaagaccc 2520
tctgtgcaga agattcaagg tacatcacgc agttcagcaa gctattgagc aaaatctgga 2580
ttcaattata ctgatttttc tccagaatat tccagattat aaactaaacc atgcactctg 2640
tttgcgaaga ggaatgttta aatctcattg catcttgaac tggccagttc agaaagaacg 2700
gataaatgcc tttcatcata aattgcaagt agcacttgga tctcggaatt cagcacatta 2760
aactcatttg aagatttgga gtcggtaaag ggatagatcc aatttataaa ggtccatcat 2820
gaatctaagt tttacttgaa agttttgtat atttatttat atgtatagat gatgatatta 2880
catcacaatc caatctcagt tttgaaatat ttcggcttat ttcattgaca tctggtttat 2940
tcactccaaa taaacacatg ggcagttaaa aacatcctct attaatagat tacccattaa 3000
ttcttgaggt gtatcacagc tttaaagggt tttaaatatt tttatataaa taagactgag 3060
agttttataa atgtaatttt ttaaaactcg agtcttactg tgtagctcag aaaggcctgg 3120
aaattaatat attagagagt catgtcttga acttatttat ctctgcctcc ctctgtctcc 3180
agagtgttgc ttttaagggc atgtagcacc acacccagct atgtacgtgt gggattttat 3240
```

```
aatgctcatt tttgagacgt ttatagaata aaagataatt gcttttatgg tataaggcta 3300
cttgaggtaa                                                        3310
```

Table 5. Amino Acid Sequence for Murine TLR3

(GenBank Accession No. AAK26117; SEQ ID NO:4)

```
MKGCSSYLMY SFGGLLSLWI LLVSSTNQCT VRYNVADCSH LKLTHIPDDL PSNITVLNLT   60
HNQLRRLPPT NFTRYSQLAI LDAGFNSISK LEPELCQILP LLKVLNLQHN ELSQISDQTF  120
VFCTNLTELD LMSNSIHKIK SNPFKNQKNL IKLDLSHNGL SSTKLGTGVQ LENLQELLLA  180
KNKILALRSE ELEFLGNSSL RKLDLSSNPL KEFSPGCFQT IGKLFALLLN NAQLNPHLTE  240
KLCWELSNTS IQNLSLANNQ LLATSESTFS GLKWTNLTQL DLSYNNLHDV GNGSFSYLPS  300
LRYLSLEYNN IQRLSPRSFY GLSNLRYLSL KRAFTKQSVS LASHPNIDDF SFQWLKYLEY  360
LNMDDNNIPS TKSNTFTGLV SLKYLSLSKT FTSLQTLTNE TFVSLAHSPL LTLNLTKNHI  420
SKIANGTFSW LGQLRILDLG LNEIEQKLSG QEWRGLRNIF EIYLSYNKYL QLSTSSFALV  480
PSLQRLMLRR VALKNVDISP SPFRPLRNLT ILDLSNNNIA NINEDLLEGL ENLEILDFQH  540
NNLARLWKRA NPGGPVNFLK GLSHLHILNL ESNGLDEIPV GVFKNLFELK SINLGLNNLN  600
KLEPFIFDDQ TSLRSLNLQK NLITSVEKDV FGPPFQNLNS LDMRFNPFDC TCESISWFVN  660
WINQTHTNIF ELSTHYLCNT PHHYYGFPLK LFDTSSCKDS APFELLFIIS TSMLLVFILV  720
VLLIHIEGWR ISFYWNVSVH RILGFKEIDT QAEQFEYTAY IIHAHKDRDW VWEHFSPMEE  780
QDQSLKFCLE ERDFEAGVLG LEAIVNSIKR SRKIIFVITH HLLKDPLCRR FKVHHAVQQA  840
IEQNLDSIIL IFLQNIPDYK LNHALCLRRG MFKSHCILNW PVQKERINAF HHKLQVALGS  900
RNSAH                                                              905
```

**Example 2. Method of Making IFN-α4 Reporter Vector**

A number of reporter vectors may be used in the practice of the invention. Some of the reporter vectors are commercially available, e.g., the luciferase reporter vectors pNF-κB-Luc (Stratagene) and pAP1-Luc (Stratagene). These two reporter vectors place the luciferase gene under control of an upstream (5') promoter region derived from genomic DNA for NF-κB or AP1, respectively. Other reporter vectors can be constructed following standard methods using the desired promoter and a vector containing a suitable reporter, such as luciferase, β-galactosidase (β-gal), chloramphenicol acetyltransferase (CAT), and other reporters known by those skilled in the art. Following are some examples of reporter vectors constructed for use in the present invention.

IFN-α4 is an immediate-early type 1 IFN. Sequence-specific PCR products for the –620 to +50 promoter region of IFN-α4 were derived from genomic DNA of human 293 cells and cloned into *Sma*I site of the pGL3-Basic Vector (Promega). The resulting expression vector includes a luciferase gene under control of an upstream (5') –620 to +50 promoter region of IFN-α4. The sequence of the –620 to +50 promoter region of IFN-α4 is provided as SEQ ID NO:11 in Table 6.

Table 6. Nucleotide Sequence of the –620 to +50 Promoter Region of Human IFN-α4 (SEQ ID NO:11)

```
agaaaaattt taaaaaatta ttcattcata tttttaggag ttttgaatga ttggatatgt   60
aattatattc atattattaa tgtgtatcta tatagatttt tattttgcat atgtactttg  120
atacaaaatt tacatgaaca aattacacta aaagttattc cacaaatata cttatcaaat  180
taagttaaat gtcaatagct tttaaactta aattttagtt taacttttct gtcattcttt  240
actttgaata aaaagagcaa actttgtagt ttttatctgt gaagtagagg tatacgtaat  300
atacataaat agatatgcca aatctgtgtt attaaaattt catgaagatt tcaattagaa  360
aaaaatacca taaaaggctt tgagtgcagg tgaaaaatag gcaatgatga aaaaaaatga  420
aaaacttttt aaacacatgt agagagtgcg taaagaaagc aaaaacagag atagaaagta  480
caactaggga atttagaaaa tggaaattag tatgttcact atttaagacc tatgcacaga  540
gcaaagtctt cagaaaacct agaggccgaa gttcaaggtt atccatctca agtagcctag  600
caatatttgc aacatcccaa tggccctgtc cttttcttta ctgatggccg tgctggtgct  660
cagctacaaa                                                        670
```

**Example 3. Method of Making IFN-α1 Reporter Vector**

IFN-α1 is a late type 1 IFN. Sequence-specific PCR products for the –140 to +9 promoter region of IFN-α1 were derived from genomic DNA of human 293 cells and cloned into *Sma*I site of the pGL3-Basic Vector (Promega). The resulting expression vector includes a luciferase gene under control of an upstream (5') –140 to +9 promoter region of IFN-α1.

**Example 4. Method of Making IFN-β Reporter Vector**

IFN-β is an immediate-early type 1 IFN. The –280 to +20 promoter region of IFN-β was derived from the pUCβ26 vector (Algarté M et al. (1999) *J Virol* 73(4):2694-702) by restriction at *EcoR*I and *Taq*I sites. The 300 bp restriction fragment was filled in by Klenow enzyme and cloned into *Nhe*I-digested and filled in pGL3-Basic Vector (Promega). The resulting expression vector includes a luciferase gene under control of an upstream (5') –280 to +20 promoter region of IFN-β. The sequence of the –280 to +20 promoter region of IFN-β is provided as SEQ ID NO:12 in Table 7.

Table 7. Nucleotide Sequence of the –280 to +20 Promoter Region of Human IFN-β (SEQ ID NO:12)

```
ttctcaggtc gtttgctttc ctttgctttc tcccaagtct tgttttacaa tttgctttag   60
tcattcactg aaactttaaa aaacattaga aaacctcaca gtttgtaaat ctttttccct  120
attatatata tcataagata ggagcttaaa taaagagttt tagaaactac taaaatgtaa  180
atgacatagg aaaactgaaa gggagaagtg aaagtgggaa attcctctga atagagagag  240
gaccatctca tataaatagg ccatacccac ggagaaagga cattctaact gcaacctttc  300
```

**Example 5. Method of Making RANTES Reporter Vector**

Transcription of the chemokine RANTES is believed to be regulated at least in part by IRF3 and by NF-κB. Lin R et al. (1999) *J Mol Cell Biol* 19(2):959-66; Genin P et al. (2000) *J Immunol* 164:5352-61. A 483 bp sequence-specific PCR product including the –397 to +5 promoter region of RANTES was derived from genomic DNA of human 293 cells, restricted with PstI and cloned into pCAT-Basic Vector (Promega) using HindIII (filled in with Klenow) and PstI sites (filled in). The –397 to +5 promoter region of RANTES was then isolated from the resulting RANTES/chloramphenicol acetyltransferase (CAT) reporter plasmid by restriction with BglII and SalI, filled in with Klenow enzyme, and cloned into the NheI site (filled in with Klenow) of the pGL3-Basic Vector (Promega). The resulting expression vector includes a luciferase gene under control of an upstream (5') –397 to +5 promoter region of RANTES. Comparison of the insert sequence –397 to +5 of Genin P et al. (2000) *J Immunol* 164:5352-61 and GenBank accession no. AB023652 (SEQ ID NO:13) revealed two point deletions (at positions 105 and 273 of SEQ ID NO:13) which do not create new restriction sites. The sequence of the –397 to +5 promoter region of RANTES is provided as SEQ ID NO:14 in Table 8.

Table 8. Nucleotide Sequence of the –397 to +5 Promoter Region of Human RANTES (SEQ ID NO:14)

```
gatctgtaat gaataagcag gaactttgaa gactcagtga ctcagtgagt aataaagact   60
cagtgacttc tgatcctgtc ctaactgcca ctccttgttg tcccaagaaa gcggcttcct  120
gctctctgag gaggacccct tccctggaag gtaaaactaa ggatgtcagc agagaaattt  180
ttccaccatt ggtgcttggt caaagaggaa actgatgagc tcactctaga tgagagagca  240
gtgagggaga gacagagact cgaatttccg gagctatttc agttttcttt tccgttttgt  300
gcaatttcac ttatgatacc ggccaatgct tggttgctat tttggaaact ccccttaggg  360
gatgcccctc aactggccct ataaagggcc agcctgagct g                      401
```

Table 9. Nucleotide Sequence of GenBank Accession No. AB023652 (SEQ ID NO:13)

```
agaaggcctt acagtgagat gggatcccag tatttattga gtttcctcat tcataaaatg   60
gggataataa tagtaaatga gttgacacgc gctaagacag tggaatagtg gctggcacag  120
ataagccctc ggtaaatggt agccaataat gatagagtat gctgtaagat atctttctct  180
ccctctgctt ctcaacaagt ctctaatcaa ttattccact ttataaacaa ggaaatagaa  240
ctcaaagaca ttaagcactt ttcccaaagg tcgcttagca agtaaatggg agagacccta  300
tgaccaggat gaaagcaaga aattcccaca agaggactca ttccaactca tatcttgtga  360
aaaggttccc aatgcccagc tcagatcaac tgcctcaatt tacagtgtga gtgtgctcac  420
ctcctttggg gactgtatat ccagaggacc ctcctcaata aaacacttta taaataacat  480
ccttccatgg atgagggaaa ggaggtaaga tctgtaatga ataagcagga actttgaaga  540
ctcagtgact cagtgagtaa taaagactca gtgacttctg atcctgtcct aactgccact  600
ccttgttgtc cccaagaaag cggcttcctg ctctctgagg aggacccctt ccctggaagg  660
taaaactaag gatgtcagca gagaaatttt tccaccattg gtgcttggtc aaagaggaaa  720
```

```
ctgatgagct cactctagat gagagagcag tgagggagag acagagactc gaatttccgg  780
aggctatttc agttttcttt tccgttttgt gcaatttcac ttatgatacc ggccaatgct  840
tggttgctat tttggaaact ccccttaggg gatgcccctc aactggccct ataaagggcc  900
agcctgagct gcagaggatt cctgcagagg atcaagacag cacgtggacc tcgcacagcc  960
tctcccacag gtaccatgaa ggtctccgcg gcagccctcg ctgtcatcct cattgctact 1020
gccctctgcg c                                                      1031
```

**Example 6. Method of Making Human IL-12 p40 Reporter Vectors**

Reporter constructs have been made using truncated (-250 to +30) and full length (-860 to +30) promoter regions derived from human IL-12 p40 genomic DNA. In one reporter construct the truncated IL-12 p40 promoter was cloned as a KpnI-XhoI insert into pβgal-Basic (Promega). The resulting expression vector includes a β gal gene under control of an upstream (5') –250 to +30 promoter region of human IL-12 p40. In a second reporter construct the full length IL-12 p40 promoter was cloned as a KpnI-XhoI insert into pβgal-Basic (Promega). The resulting expression vector includes a β gal gene under control of an upstream (5') –860 to +30 promoter region of human IL-12 p40. In a third reporter construct the truncated –250 to +30 promoter region of human IL-12 p40 was cloned into the pGL3-Basic Vector (Promega). The resulting expression vector includes a luciferase gene under control of an upstream (5') –250 to +30 promoter region of human IL-12 p40. In a fourth reporter construct the full length IL-12 p40 promoter of human IL-12 p40 was cloned into the pGL3-Basic Vector (Promega). The resulting expression vector includes a luciferase gene under control of an upstream (5') –860 to +30 promoter region of human IL-12 p40.

**Example 7. Method of Making Human IL-6 Reporter Vectors**

Reporter constructs are made using the -235 to +7 promoter region derived from human IL-6 genomic DNA. In one reporter construct the IL-6 promoter region is cloned as a KpnI-XhoI insert into pGL3-Basic Vector (Promega). The resulting expression vector includes a luciferase gene under control of an upstream (5') -235 to +7 promoter region derived from human IL-6 genomic DNA.

**Example 8. Method of Making Human IL-8 Reporter Vectors**

Reporter constructs have been made using a -546 to +44 and a truncated -133 to +44 promoter region derived from human IL-8 genomic DNA. Mukaida N et al. (1989) *J Immunol* 143:1366-71. In each reporter construct the IL-8 promoter region

was cloned as a KpnI-XhoI insert into pGL3-Basic Vector (Promega). One of the resulting expression vectors includes a luciferase gene under control of an upstream (5') -546 to +44 promoter region derived from human IL-8 genomic DNA. Another of the resulting expression vectors includes a luciferase gene under control of an upstream (5') -133 to +44 promoter region derived from human IL-8 genomic DNA.

**Example 9. Sequence Comparison of Human TLR3 and Human TLR9**

Human TLR3 and TLR9 are homologous proteins with several structural commonalities. Both appear to be transmembrane proteins with an extracellular domain and an intracellular domain. Common characteristics include a signal sequence and transmembranal domain. Similarities common to most TLRs include a cysteine rich domain and a TIR domain. Most TLRs have leucine rich repeats (LRR) in their extracellular domain. TLR3, TLR7, TLR8, and TLR9 appear to have similar structures. The regularity of the leucine repeats are shown below for TLR3 and TLR9. These four TLRs can be broken into two extracellular subdomains, domain 1 and 2, by virtue of a separation by an unstructured hinge region. TLR7, TLR8, and TLR9 have 14 LRR in domain 1 and 12 LRR in domain 2. TLR9 is a known nucleic acid binder, interacting with CpG-DNA. It has been suspected that TLR7 and TLR8 most likely also interact with nucleic acids. TLR3 has a similar 11 LRR in domain 1 and has 12 LRR in domain 2, lacking the initial 3 repeats common to TLR7, TLR8, and TLR9. Based on structural consideration it is hypothesized that TLR3 interacts with nucleic acids or similar structures.

The structure of TLR3 differs from TLR7, TLR8, and TLR9 in an interesting character. Referring to Table 13, within the TIR domain it has been shown that a proline (shown in bold) is required for MyD88 interaction. MyD88 is required for TLR9 to transduce signal for the activation of NF-κB. Both TLR7 and TLR8 also have this proline. TLR3 however has an alanine at this position (also shown in bold). It is believed by the applicant that this difference may disallow MyD88 interaction with TLR3 and thus result in an altered signal transduction pattern compared to, e.g., TLR9.

Table 10. Sequence Alignment of hTLR9 (SEQ ID NO:6) and hTLR3 (SEQ ID NO:2)

SIGNAL SEQUENCE

| | | |
|---|---|---|
| hTLR9 | MGFCRSALHPLSLLVQAIMLAMTLALGTLPAFLPCELQPHGLVNCNW | 47 |
| hTLR3 | MRQTLPCIYFWGGLLPFGMLCASSTTKCTVSHEVADC | 37 |

DOMAIN 1 LEUCINE RICH REPEATS

| | | |
|---|---|---|
| hTLR9 | LFLKSVPHFSMAAPRGNVTSLSLSSN | 73 |
| hTLR9 | RIHHLHDSDFAHLPSLRHLNLKWN | 97 |
| hTLR9 | CPPVGLSPMHFPCHMTIEPSTFLAVPTLEELNLSYN | 133 |
| hTLR9 | NIMTVPALPKSLISLSLSHT | 153 |
| hTLR3 | SHLKLTQVPDDLPTNITVLNLTHN | 61 |
| hTLR9 | NILMLDSASLAGLHALRFLFMDGN | 177 |
| hTLR3 | QLRRLPAANFTRYSQLTSLDVGFN | 85 |
| hTLR9 | CYYKNPCRQALEVAPGALLGLGNLTHLSLKYN | 209 |
| hTLR3 | TISKLEPELCQKLPMLKVLNLQHN | 109 |
| hTLR9 | NLTVVPRNLPSSLEYLLLSYN | 230 |
| hTLR3 | ELSQLSDKTFAFCTNLTELHLMSN | 133 |
| hTLR9 | RIVKLAPEDLANLTALRVLDVGGN | 254 |
| hTLR3 | SIQKIKNNPFVKQKNLITLDLSHN | 157 |
| hTLR9 | *CRRCDHAPNPCMECPRHFPQ*LHPDTFSHLSRLEGLVLKDS | 294 |
| hTLR3 | GLSSTKLGTQVQLENLQELLLSNN | 181 |
| hTLR9 | SLSWLNASWFRGLGNLRVLDLSEN | 318 |
| hTLR3 | KIQALKSEELDIFANSSLKKLELSSN | 207 |
| hTLR9 | FLYKCITKTKAFQGLTQLRKLNLSFN | 344 |
| hTLR3 | QIKEFSPGCFHAIGRLFGLFLNNV | 231 |
| hTLR9 | YQKRVSFAHLSLAPSFGSLVALKELDMHGI | 374 |
| hTLR3 | QLGPSLTEKLCLELANTSIRNLSLSNS | 258 |
| hTLR9 | FFRSLDETTLRPLARLPMLQTLRLQMN | 401 |
| hTLR3 | QLSTTSNTTFLGLKWTNLTMLDLSYN | 284 |
| hTLR9 | FINQAQLGIFRAFPGLRYVDLSDN | 425 |
| hTLR3 | NLNVVGNDSFAWLPQLEYFFLEYN | 308 |

HINGE REGION

| | | |
|---|---|---|
| hTLR9 | RISGASELTATMGEADGGEKVWLQPGDLAPAPV | 458 |
| hTLR3 | NIQHLFSHSLHGLFNVRYLNLKRSFTKQSISLA | 341 |

DOMAIN 2 LEUCINE RICH REPEATS

| | | |
|---|---|---|
| hTLR9 | DTPSSEDFRPNCSTLNFTLDLSRN | 482 |
| hTLR3 | SLPKIDDFSFQWLKCLEHLNMEDN | 365 |
| hTLR9 | NLVTVQPEMFAQLSHLQCLRLSHN | 506 |
| hTLR3 | DIPGIKSNMFTGLINLKYLSLSNS | 389 |

| | | | |
|---|---|---|---|
| | hTLR9 | CISQAVNGSQFLPLTGLQVLDLSHN | 531 |
| | hTLR3 | FTSLRTLTNETFVSLAHSPLHILNLTKN | 417 |
| | hTLR9 | KLDLYHEHSFTELPRLEALDLSYN | 555 |
| | hTLR3 | KISKIESDAFSWLGHLEVLDLGLN | 441 |
| | hTLR9 | SQPFGMQGVGHNFSFVAHLRTLRHLSLAHN | 585 |
| | hTLR3 | EIGQELTGQEWRGLENIFEIYLSYN | 466 |
| | hTLR9 | NIHSQVSQQLCSTSLRALDFSGN | 608 |
| | hTLR3 | KYLQLTRNSFALVPSLQRLMLRRV | 490 |
| | hTLR9 | ALGHMWAEGDLYLHFFQGLSGLIWLDLSQN | 638 |
| | hTLR3 | ALKNVDSSPSPFQPLRNLTILDLSNN | 516 |
| | hTLR9 | RLHTLLPQTLRNLPKSLQVLRLRDN | 663 |
| | hTLR3 | NIANINDDMLEGLEKLEILDLQHN | 540 |
| | hTLR9 | YLAFFKWWSLHFLPKLEVLDLAGN | 687 |
| | hTLR3 | NLARLWKHANPGGPIYFLKGLSHLHILNLESN | 572 |
| | hTLR9 | QLKALTNGSLPAGTRLRRLDVSCN | 711 |
| | hTLR3 | GFDEIPVEVFKDLFELKIIDLGLN | 596 |
| | hTLR9 | SISFVAPGFFSKAKELRELNLSAN | 735 |
| | hTLR3 | NLNTLPASVFNNQVSLKSLNLQKN | 620 |
| | hTLR9 | ALKTVDHSWFGPLASALQILDVSAN | 760 |
| | hTLR3 | LITSVEKKVFGPAFRNLTELDMRFN | 645 |

CYSTEINE RICH DOMAIN

| | | |
|---|---|---|
| hTLR9 | PLHCACG**AAFMDFLLEVQAAVPGLPSRVKCGSPGQLQGLSIFAQD | 805 |
| hTLR3 | PFDCTCESIAWFVNWINETHTNIPELSSHYLCNTPPHYHGFPVRLFD | 692 |
| hTLR9 | LRLCLDEALSWDCFA | 820 |
| hTLR3 | TSSCKDSAPFELFFM | 707 |

TRANSMEMBRANAL DOMAIN

| | | |
|---|---|---|
| hTLR9 | LSLLAVALGLGVPMLHHL | 838 |
| hTLR3 | INTSILLIFIFIVLLIHF | 725 |

TIR DOMAIN

| | | |
|---|---|---|
| hTLR9 | CGWDLWYCFHLCLAWLPWRGRQSGRDEDALPYDAFVVFDKTQSAVAD | 885 |
| hTLR3 | EGWRISFYWNVSVHRVLGFKEIDRQTEQFE*YAAYIIHAYK***DKD | 768 |
| hTLR9 | WVYNELRGQLEECRGRWALRLCLEERDWLPGKTLFENLWASVYGSRK | 932 |
| hTLR3 | WVW***EHFSSMEKEDQSLKFCLEERDFEAGVFELEAIVNSIKRSRK | 812 |
| hTLR9 | TLFVLAHTD*RVSGLLRASFLLAQQRLLEDRKDVVVLVILSPDGRRS | 978 |
| hTLR3 | IIFVITHHLLKDPLCKRFKVHHAVQQAIEQNLDSIILVFLEEIPDYK | 859 |
| hTLR9 | ***RYVRLRQRLCRQSVLLWPHQPSGQRSFWAQLGMALTRDNHHFYN | 1022 |
| hTLR3 | LNHALCLRRGMFKSHCILNWPVQKERIGAFRHKLQVALGSKNSVH | 904 |
| hTLR9 | RNFCQGPTAE | 1032 |

**Example 10. Reconstitution of TLR9 Signaling in 293 Fibroblasts**

Methods for cloning murine and human TLR9 have been described in pending U.S. Patent Application No. 09/954,987 and corresponding published PCT application PCT/US01/29229, both filed September 17, 2001, the contents of which are incorporated by reference. Human TLR9 cDNA and murine TLR9 cDNA in pT-Adv vector (from Clonetech) were individually cloned into the expression vector pcDNA3.1(-) from Invitrogen using the EcoRI site. Utilizing a "gain of function" assay it was possible to reconstitute human TLR9 (hTLR9) and murine TLR9 (mTLR9) signaling in CpG-DNA non-responsive human 293 fibroblasts (ATCC, CRL-1573). The expression vectors mentioned above were transfected into 293 fibroblast cells using the calcium phosphate method.

Table 11. cDNA Sequence for Human TLR9

(GenBank Accession No. AF245704; SEQ ID NO:5)

```
aggctggtat aaaaatctta cttcctctat tctctgagcc gctgctgccc ctgtgggaag   60
ggacctcgag tgtgaagcat ccttccctgt agctgctgtc cagtctgccc gccagaccct  120
ctggagaagc ccctgccccc cagcatgggt ttctgccgca gcgccctgca cccgctgtct  180
ctcctggtgc aggccatcat gctggccatg accctggccc tgggtacctt gcctgccttc  240
ctaccctgtg agctccagcc ccacggcctg gtgaactgca actggctgtt cctgaagtct  300
gtgccccact tctccatggc agcaccccgt ggcaatgtca ccagcctttc cttgtcctcc  360
aaccgcatcc accacctcca tgattctgac tttgcccacc tgcccagcct gcggcatctc  420
aacctcaagt ggaactgccc gccggttggc ctcagcccca tgcacttccc ctgccacatg  480
accatcgagc ccagcacctt cttggctgtg cccaccctgg aagagctaaa cctgagctac  540
aacaacatca tgactgtgcc tgcgctgccc aaatccctca tatccctgtc cctcagccat  600
accaacatcc tgatgctaga ctctgccagc ctcgccggcc tgcatgccct gcgcttccta  660
ttcatggacg gcaactgtta ttacaagaac ccctgcaggc aggcactgga ggtggccccg  720
ggtgccctcc ttggcctggg caacctcacc cacctgtcac tcaagtacaa caacctcact  780
gtggtgcccc gcaacctgcc ttccagcctg gagtatctgc tgttgtccta caaccgcatc  840
gtcaaactgg cgcctgagga cctggccaat ctgaccgccc tgcgtgtgct cgatgtgggc  900
ggaaattgcc gccgctgcga ccacgctccc aacccctgca tggagtgccc tcgtcacttc  960
ccccagctac atcccgatac cttcagccac ctgagccgtc ttgaaggcct ggtgttgaag 1020
gacagttctc tctcctggct gaatgccagt tggttccgtg ggctgggaaa cctccgagtg 1080
ctggacctga gtgagaactt cctctacaaa tgcatcacta aaaccaaggc cttccagggc 1140
ctaacacagc tgcgcaagct taacctgtcc ttcaattacc aaaagagggt gtcctttgcc 1200
cacctgtctc tggccccttc cttcgggagc ctggtcgccc tgaaggagct ggacatgcac 1260
ggcatcttct tccgctcact cgatgagacc acgctccggc cactggcccg cctgcccatg 1320
ctccagactc tgcgtctgca gatgaacttc atcaaccagg cccagctcgg catcttcagg 1380
gccttccctg gcctgcgcta cgtggacctg tcggacaacc gcatcagcgg agcttcggag 1440
ctgacagcca ccatggggga ggcagatgga ggggagaagg tctggctgca gcctggggac 1500
cttgctccgg ccccagtgga cactcccagc tctgaagact tcaggcccaa ctgcagcacc 1560
ctcaacttca ccttggatct gtcacggaac aacctggtga ccgtgcagcc ggagatgttt 1620
gcccagctct cgcacctgca gtgcctgcgc ctgagccaca actgcatctc gcaggcagtc 1680
aatggctccc agttcctgcc gctgaccggt ctgcaggtgc tagacctgtc ccgcaataag 1740
ctggacctct accacgagca ctcattcacg gagctaccgc gactggaggc cctggacctc 1800
agctacaaca gccagccctt tggcatgcag ggcgtgggcc acaacttcag cttcgtggct 1860
cacctgcgca ccctgcgcca cctcagcctg gcccacaaca acatccacag ccaagtgtcc 1920
```

```
cagcagctct gcagtacgtc gctgcgggcc ctggacttca gcggcaatgc actgggccat 1980
atgtgggccg agggagacct ctatctgcac ttcttccaag gcctgagcgg tttgatctgg 2040
ctggacttgt cccagaaccg cctgcacacc ctcctgcccc aaaccctgcg caacctcccc 2100
aagagcctac aggtgctgcg tctccgtgac aattacctgg ccttctttaa gtggtggagc 2160
ctccacttcc tgcccaaact ggaagtcctc gacctggcag gaaaccggct gaaggccctg 2220
accaatggca gcctgcctgc tggcacccgg ctccggaggc tggatgtcag ctgcaacagc 2280
atcagcttcg tggcccccgg cttcttttcc aaggccaagg agctgcgaga gctcaacctt 2340
agcgccaacg ccctcaagac agtggaccac tcctggtttg ggcccctggc gagtgccctg 2400
caaatactag atgtaagcgc caaccctctg cactgcgcct gtggggcggc ctttatggac 2460
ttcctgctgg aggtgcaggc tgccgtgccc ggtctgccca gccgggtgaa gtgtggcagt 2520
ccgggccagc tccagggcct cagcatcttt gcacaggacc tgcgcctctg cctggatgag 2580
gccctctcct gggactgttt cgccctctcg ctgctggctg tggctctggg cctgggtgtg 2640
cccatgctgc atcacctctg tggctgggac ctctggtact gcttccacct gtgcctggcc 2700
tggcttccct ggcgggggcg gcaaagtggg cgagatgagg atgccctgcc ctacgatgcc 2760
ttcgtggtct tcgacaaaac gcagagcgca gtggcagact gggtgtacaa cgagcttcgg 2820
gggcagctgg aggagtgccg tgggcgctgg gcactccgcc tgtgcctgga ggaacgcgac 2880
tggctgcctg gcaaaaccct ctttgagaac ctgtgggcct cggtctatgg cagccgcaag 2940
acgctgtttg tgctggccca cacggaccgg gtcagtggtc tcttgcgcgc cagcttcctg 3000
ctggcccagc agcgcctgct ggaggaccgc aaggacgtcg tggtgctggt gatcctgagc 3060
cctgacggcc gccgctcccg ctacgtgcgg ctgcgccagc gcctctgccg ccagagtgtc 3120
ctcctctggc cccaccagcc cagtggtcag cgcagcttct gggcccagct gggcatggcc 3180
ctgaccaggg acaaccacca cttctataac cggaacttct gccagggacc cacggccgaa 3240
tagccgtgag ccggaatcct gcacggtgcc acctccacac tcacctcacc tctgcctgcc 3300
tggtctgacc ctcccctgct cgcctccctc accccacacc tgacacagag ca          3352
```

Table 12. Amino Acid Sequence for Human TLR9

(GenBank Accession No. AAF78037; SEQ ID NO:6)

```
MGFCRSALHP LSLLVQAIML AMTLALGTLP AFLPCELQPH GLVNCNWLFL KSVPHFSMAA   60
PRGNVTSLSL SSNRIHHLHD SDFAHLPSLR HLNLKWNCPP VGLSPMHFPC HMTIEPSTFL  120
AVPTLEELNL SYNNIMTVPA LPKSLISLSL SHTNILMLDS ASLAGLHALR FLFMDGNCYY  180
KNPCRQALEV APGALLGLGN LTHLSLKYNN LTVVPRNLPS SLEYLLLSYN RIVKLAPEDL  240
ANLTALRVLD VGGNCRRCDH APNPCMECPR HFPQLHPDTF SHLSRLEGLV LKDSSLSWLN  300
ASWFRGLGNL RVLDLSENFL YKCITKTKAF QGLTQLRKLN LSFNYQKRVS FAHLSLAPSF  360
GSLVALKELD MHGIFFRSLD ETTLRPLARL PMLQTLRLQM NFINQAQLGI FRAFPGLRYV  420
DLSDNRISGA SELTATMGEA DGGEKVWLQP GDLAPAPVDT PSSEDFRPNC STLNFTLDLS  480
RNNLVTVQPE MFAQLSHLQC LRLSHNCISQ AVNGSQFLPL TGLQVLDLSR NKLDLYHEHS  540
FTELPRLEAL DLSYNSQPFG MQGVGHNFSF VAHLRTLRHL SLAHNNIHSQ VSQQLCSTSL  600
RALDFSGNAL GHMWAEGDLY LHFFQGLSGL IWLDLSQNRL HTLLPQTLRN LPKSLQVLRL  660
RDNYLAFFKW WSLHFLPKLE VLDLAGNRLK ALTNGSLPAG TRLRRLDVSC NSISFVAPGF  720
FSKAKELREL NLSANALKTV DHSWFGPLAS ALQILDVSAN PLHCACGAAF MDFLLEVQAA  780
VPGLPSRVKC GSPGQLQGLS IFAQDLRLCL DEALSWDCFA LSLLAVALGL GVPMLHHLCG  840
WDLWYCFHLC LAWLPWRGRQ SGRDEDALPY DAFVVFDKTQ SAVADWVYNE LRGQLEECRG  900
RWALRLCLEE RDWLPGKTLF ENLWASVYGS RKTLFVLAHT DRVSGLLRAS FLLAQQRLLE  960
DRKDVVVLVI LSPDGRRSRY VRLRQRLCRQ SVLLWPHQPS GQRSFWAQLG MALTRDNHHF 1020
YNRNFCQGPT AE                                                      1032
```

Table 13. cDNA Sequence for Murine TLR9

(GenBank Accession No. AF348140; SEQ ID NO:7)

```
tgtcagaggg agcctcggga gaatcctcca tctcccaaca tggttctccg tcgaaggact   60
ctgcaccct tgtccctcct ggtacaggct gcagtgctgg ctgagactct ggccctgggt   120
accctgcctg ccttcctacc ctgtgagctg aagcctcatg gcctggtgga ctgcaattgg   180
ctgttcctga agtctgtacc ccgtttctct gcggcagcat cctgctccaa catcacccgc   240
ctctccttga tctccaaccg tatccaccac ctgcacaact ccgacttcgt ccacctgtcc   300
aacctgcggc agctgaacct caagtggaac tgtccaccca ctggccttag ccccctgcac   360
```

```
ttctcttgcc acatgaccat tgagcccaga accttcctgg ctatgcgtac actggaggag  420
ctgaacctga gctataatgg tatcaccact gtgccccgac tgcccagctc cctggtgaat  480
ctgagcctga gccacaccaa catcctggtt ctagatgcta acagcctcgc cggcctatac  540
agcctgcgcg ttctcttcat ggacgggaac tgctactaca agaacccctg cacaggagcg  600
gtgaaggtga ccccaggcgc cctcctgggc ctgagcaatc tcacccatct gtctctgaag  660
tataacaacc tcacaaaggt gccccgccaa ctgcccccca gcctggagta cctcctggtg  720
tcctataacc tcattgtcaa gctggggcct gaagacctgg ccaatctgac ctcccttcga  780
gtacttgatg tgggtgggaa ttgccgtcgc tgcgaccatg cccccaatcc ctgtatagaa  840
tgtggccaaa agtccctcca cctgcaccct gagaccttcc atcacctgag ccatctggaa  900
ggcctggtgc tgaaggacag ctctctccat acactgaact cttcctggtt ccaaggtctg  960
gtcaacctct cggtgctgga cctaagcgag aactttctct atgaaagcat caaccacacc 1020
aatgcctttc agaacctaac ccgcctgcgc aagctcaacc tgtccttcaa ttaccgcaag 1080
aaggtatcct ttgcccgcct ccacctggca agttccttca agaacctggt gtcactgcag 1140
gagctgaaca tgaacggcat cttcttccgc tcgctcaaca agtacacgct cagatggctg 1200
gccgatctgc ccaaactcca cactctgcat cttcaaatga acttcatcaa ccaggcacag 1260
ctcagcatct ttggtacctt ccgagcccct cgctttgtgg acttgtcaga caatcgcatc 1320
agtgggcctt caacgctgtc agaagccacc cctgaagagg cagatgatgc agagcaggag 1380
gagctgttgt ctgcggatcc tcacccagct ccactgagca cccctgcttc taagaacttc 1440
atggacaggt gtaagaactt caagttcacc atggacctgt ctcggaacaa cctggtgact 1500
atcaagccag agatgtttgt caatctctca cgcctccagt gtcttagcct gagccacaac 1560
tccattgcac aggctgtcaa tggctctcag ttcctgccgc tgactaatct gcaggtgctg 1620
gacctgtccc ataacaaact ggacttgtac cactggaaat cgttcagtga gctaccacag 1680
ttgcaggccc tggacctgag ctacaacagc cagcccttta gcatgaaggg tataggccac 1740
aatttcagtt ttgtggccca tctgtccatg ctacacagcc ttagcctggc acacaatgac 1800
attcataccc gtgtgtcctc acatctcaac agcaactcag tgaggtttct tgacttcagc 1860
ggcaacggta tgggccgcat gtgggatgag gggggccttt atctccattt cttccaaggc 1920
ctgagtggcc tgctgaagct ggacctgtct caaaataacc tgcatatcct ccggccccag 1980
aaccttgaca acctccccaa gagcctgaag ctgctgagcc tccgagacaa ctacctatct 2040
ttctttaact ggaccagtct gtccttcctg cccaacctgg aagtcctaga cctggcaggc 2100
aaccagctaa aggccctgac caatggcacc ctgcctaatg gcaccctcct ccagaaactg 2160
gatgtcagca gcaacagtat cgtctctgtg gtcccagcct tcttcgctct ggcggtcgag 2220
ctgaaagagg tcaacctcag ccacaacatt ctcaagacgg tggatcgctc ctggtttggg 2280
cccattgtga tgaacctgac agttctagac gtgagaagca accctctgca ctgtgcctgt 2340
ggggcagcct tcgtagactt actgttggag gtgcagacca aggtgcctgg cctggctaat 2400
ggtgtgaagt gtggcagccc cggccagctg cagggccgta gcatcttcgc acaggacctg 2460
cggctgtgcc tggatgaggt cctctcttgg gactgctttg gcctttcact cttggctgtg 2520
gccgtgggca tggtggtgcc tatactgcac catctctgcg gctgggacgt ctggtactgt 2580
tttcatctgt gcctggcatg gctacctttg ctggcccgca gccgacgcag cgcccaagct 2640
ctcccctatg atgccttcgt ggtgttcgat aaggcacaga gcgcagttgc ggactgggtg 2700
tataacgagc tgcgggtgcg gctggaggag cggcgcggtc gccgagccct acgcttgtgt 2760
ctggaggacc gagattggct gcctggccag acgctcttcg agaacctctg ggcttccatc 2820
tatgggagcc gcaagactct atttgtgctg gcccacacgg accgcgtcag tggcctcctg 2880
cgcaccagct tcctgctggc tcagcagcgc ctgttggaag accgcaagga cgtggtggtg 2940
ttggtgatcc tgcgtccgga tgcccaccgc tcccgctatg tgcgactgcg ccagcgtctc 3000
tgccgccaga gtgtgctctt ctggccccag cagcccaacg ggcaggggg cttctgggcc 3060
cagctgagta cagccctgac tagggacaac cgccacttct ataaccagaa cttctgccgg 3120
ggacctacag cagaatagct cagagcaaca gctggaaaca gctgcatctt catgcctggt 3180
tcccgagttg ctctgcctgc                                           3200
```

Table 14. Amino Acid Sequence for Murine TLR9

(GenBank Accession No. AAK29625; SEQ ID NO:8)

```
MVLRRRTLHP LSLLVQAAVL AETLALGTLP AFLPCELKPH GLVDCNWLFL KSVPRFSAAA   60
SCSNITRLSL ISNRIHHLHN SDFVHLSNLR QLNLKWNCPP TGLSPLHFSC HMTIEPRTFL  120
AMRTLEELNL SYNGITTVPR LPSSLVNLSL SHTNILVLDA NSLAGLYSLR VLFMDGNCYY  180
KNPCTGAVKV TPGALLGLSN LTHLSLKYNN LTKVPRQLPP SLEYLLVSYN LIVKLGPEDL  240
ANLTSLRVLD VGGNCRRCDH APNPCIECGQ KSLHLHPETF HHLSHLEGLV LKDSSLHTLN  300
SSWFQGLVNL SVLDLSENFL YESINHTNAF QNLTRLRKLN LSFNYRKKVS FARLHLASSF  360
```

```
KNLVSLQELN MNGIFFRSLN KYTLRWLADL PKLHTLHLQM NFINQAQLSI FGTFRALRFV  420
DLSDNRISGP STLSEATPEE ADDAEQEELL SADPHPAPLS TPASKNFMDR CKNFKFTMDL  480
SRNNLVTIKP EMFVNLSRLQ CLSLSHNSIA QAVNGSQFLP LTNLQVLDLS HNKLDLYHWK  540
SFSELPQLQA LDLSYNSQPF SMKGIGHNFS FVAHLSMLHS LSLAHNDIHT RVSSHLNSNS  600
VRFLDFSGNG MGRMWDEGGL YLHFFQGLSG LLKLDLSQNN LHILRPQNLD NLPKSLKLLS  660
LRDNYLSFFN WTSLSFLPNL EVLDLAGNQL KALTNGTLPN GTLLQKLDVS SNSIVSVVPA  720
FFALAVELKE VNLSHNILKT VDRSWFGPIV MNLTVLDVRS NPLHCACGAA FVDLLLEVQT  780
KVPGLANGVK CGSPGQLQGR SIFAQDLRLC LDEVLSWDCF GLSLLAVAVG MVVPILHHLC  840
GWDVWYCFHL CLAWLPLLAR SRRSAQALPY DAFVVFDKAQ SAVADWVYNE LRVRLEERRG  900
RRALRLCLED RDWLPGQTLF ENLWASIYGS RKTLFVLAHT DRVSGLLRTS FLLAQQRLLE  960
DRKDVVVLVI LRPDAHRSRY VRLRQRLCRQ SVLFWPQQPN GQGGFWAQLS TALTRDNRHF 1020
YNQNFCRGPT AE                                                     1032
```

Since NF-κB activation is central to the IL-1/TLR signal transduction pathway (Medzhitov R et al. (1998) *Mol Cell* 2:253-258 (1998); Muzio M et al. (1998) *J Exp Med* 187:2097-101), cells were transfected with hTLR9 or co-transfected with hTLR9 and an NF-κB-driven luciferase reporter construct. Human 293 fibroblast cells were transiently transfected with (**Figure 1A**) hTLR9 and a six-times NF-κB-luciferase reporter plasmid (NF-κB-luc, kindly provided by Patrick Baeuerle, Munich, Germany) or (**Figure 1B**) with hTLR9 alone. After stimulus with CpG-ODN (2006, 2μM, TCGTCGTTTTGTCGTTTTGTCGTT, SEQ ID NO:15), GpC-ODN (2006-GC, 2μM, TGCTGCTTTTGTGCTTTTGTGCTT, SEQ ID NO:16), LPS (100 ng/ml) or media, NF-κB activation by luciferase readout (8h, **Figure 1A**) or IL-8 production by ELISA (48h, **Figure 1B**) were monitored. Results are representative of three independent experiments. **Figure 1** shows that cells expressing hTLR9 responded to CpG-DNA but not to LPS.

**Figure 2** demonstrates the same principle for the transfection of mTLR9. Human 293 fibroblast cells were transiently transfected with mTLR9 and the NF-κB-luc construct (**Figure 2**). Similar data was obtained for IL-8 production (not shown). Thus expression of TLR9 (human or mouse) in 293 cells results in a gain of function for CpG-DNA stimulation similar to hTLR4 reconstitution of LPS responses.

To generate stable clones expressing human TLR9, murine TLR9, or either TLR9 with the NF-κB-luc reporter plasmid, 293 cells were transfected in 10 cm plates ($2x10^6$ cells/plate) with 16 μg of DNA and selected with 0.7 mg/ml G418 (PAA Laboratories GmbH, Cölbe, Germany). Clones were tested for TLR9 expression by RT-PCR, for example as shown in **Figure 3**. The clones were also screened for IL-8

production or NF-κB-luciferase activity after stimulation with ODN. Four different types of clones were generated.

293-hTLR9-luc: expressing human TLR9 and 6-fold NF-κB–luciferase reporter
293-mTLR9-luc: expressing murine TLR9 and 6-fold NF-κB–luciferase reporter
293-hTLR9: expressing human TLR9
293-mTLR9: expressing murine TLR9

**Figure 4** demonstrates the responsiveness of a stable 293-hTLR9-luc clone after stimulation with CpG-ODN (2006, 2μM), GpC-ODN (2006-GC, 2μM), Me-CpG-ODN (2006 methylated, 2μM; TZGTZGTTTTGTZGTTTTGTZGTT, Z = 5-methylcytidine, SEQ ID NO:17), LPS (100 ng/ml) or media, as measured by monitoring NF-κB activation. Similar results were obtained utilizing IL-8 production with the stable clone 293-hTLR9. 293-mTLR9-luc were also stimulated with CpG-ODN (1668, 2μM; TCCATGACGTTCCTGATGCT, SEQ ID NO:18), GpC-ODN (1668-GC, 2μM; TCCATGAGCTTCCTGATGCT, SEQ ID NO:19), Me-CpG-ODN (1668 methylated, 2μM; TCCATGAZGTTCCTGATGCT, Z = 5-methylcytidine, SEQ ID NO:20), LPS (100 ng/ml) or media, as measured by monitoring NF-κB activation (**Figure 5**). Similar results were obtained utilizing IL-8 production with the stable clone 293-mTLR9. Results are representative of at least two independent experiments. These results demonstrate that CpG-DNA non-responsive cell lines can be stably genetically complemented with TLR9 to become responsive to CpG-DNA in a motif-specific manner. These cells can be used for screening of optimal ligands for innate immune responses driven by TLR9 in multiple species.

**Example 11. Reconstitution of TLR3 Signaling in 293 Fibroblasts**

Human TLR3 cDNA and murine TLR3 cDNA in pT-Adv vector (from Clonetech) were individually cloned into the expression vector pcDNA3.1(-) from Invitrogen using the EcoRI site. The resulting expression vectors mentioned above were transfected into CpG-DNA non-responsive human 293 fibroblast cells (ATCC, CRL-1573) using the calcium phosphate method. Utilizing a "gain of function" assay it

was possible to reconstitute human TLR3 (hTLR3) and murine TLR3 (mTLR3) signaling in 293 fibroblast cells.

Since NF-κB activation is central to the IL-1/TLR signal transduction pathway (Medzhitov R et al. (1998) *Mol Cell* 2:253-8; Muzio M et al. (1998) *J Exp Med* 187:2097-101), in a first set of experiments human 293 fibroblast cells were transfected with hTLR3 alone or co-transfected with hTLR3 and an NF-κB-driven luciferase reporter construct.

Likewise, in a second set of experiments, 293 fibroblast cells were transfected with hTLR3 alone or co-transfected with hTLR3 and an IFN-α4-driven luciferase reporter construct (described in Example 2 above).

In a third group of experiments, 293 fibroblast cells were transfected with hTLR3 alone or co-transfected with hTLR3 and a RANTES-driven luciferase reporter construct (described in Example 5 above).

**Example 12. Proline to Histidine Mutation P915H in the TIR Domain of Human and MurineTLR9 Alters TLR9 Signaling**

Toll-like receptors have a cytoplasmic Toll/IL-1 receptor (TIR) homology domain which initiates signaling after binding of the adapter molecule MyD88. Medzhitov R et al. (1998) *Mol Cell* 2:253-8; Kopp EB et al. (1999) *Curr Opin Immunol* 11:15-8. Reports by others have shown that a single point mutation in the signaling TIR domain in murine TLR4 (Pro712 to His, P712H) or human TLR2 (Pro681 to His, P681H) abolishes host immune response to lipopolysaccharide or gram-positive bacteria, respectively. Poltorak A et al. (1998) *Science* 282:2085-8; Underhill DM et al. (1999) *Nature* 401:811-5. Through site-specific mutagenesis the equivalent proline (P) at position 915 of human TLR9 and murine TLR9 were mutated to histidine (H; P915H). These mutations were generated by the use of the primers 5'-GCGACTGGCTGCATGGCAAAACCCTCTTTG-3' (SEQ ID NO:21) and 5'-CAAAGAGGGTTTTGCCATGCAGCCAGTCGC-3' (SEQ ID NO:22) for human TLR9 and the primers 5'-CGAGATTGGCTGCATGGCCAGACGCTCTTC-3' (SEQ ID NO:23) and 5'-GAAGAGCGTCTGGCCATGCAGCCAATCTCG-3' (SEQ ID NO:24) for murine TLR9. Expression vectors for the mutant TLR9s, hTLR9-P915H

and mTLR9-P915H, were constructed and verified using standard recombinant DNA techniques.

For the stimulation of human TLR9 variant, hTLR9-P915H, 293 cells were transiently transfected with expression vector for hTLR9 or hTLR9-P915H and stimulated after 16 hours with ODN 2006 or ODN 1668 at various concentrations. Likewise for the stimulation of murine TLR9 variant, mTLR9-P915H, 293 cells were transiently transfected with expression vector for mTLR9 or mTLR9-P915H and stimulated after 16 hours with ODN 2006 or ODN 1668 at various concentrations. After 48 hours of stimulation, supernatant was harvested and IL-8 production was measured by ELISA. Results demonstrated that TLR9 activity can be destroyed by the P915H mutation in the TIR domain of both human and murine TLR9.

**Example 13. Exchange of the TIR Domain Between Human TLR3 and Human TLR9 (hTLR3-TIR9 and hTLR9-TIR3)**

While TLR3 and TLR9 share many structural features, TLR3, by virtue of its having an alanine rather than proline at a critical position in the TIR domain, may not be able to signal via MyD88 as does TLR9. The chimeric TLRs described here can be used in the screening assays of the invention. To generate molecules consisting of human extracellular TLR3 and the TIR domain of human TLR9 (hTLR3-TIR9), the following approach can be used. Through site-specific mutagenesis a ClaI restriction site is introduced in human TLR3 and human TLR9. For human TLR9 the DNA sequence 5'-GGCCTCAGCATCTTT-3' (3026-3040, SEQ ID NO:25) is mutated to 5'-GGCCTATCGATTTTT-3' (SEQ ID NO:26), introducing a ClaI site (underlined in the sequence) but leaving the amino acid sequence (GLSIF, aa 798-802) unchanged. For human TLR3 the DNA sequence 5'-GGGTTCCCAGTGAGA-3' (2112-2126, SEQ ID NO:27) is mutated to 5'-GGGTTATCGATTAGA-3' (SEQ ID NO:28), introducing a ClaI site and creating the amino acid sequence (GLSIR, aa 685-689) which differs in three positions (aa 686, 687, 688) from the wildtype human TLR3 sequence (GFPVR, aa 685-689).

**hTLR3-TIR9**. The primers used for human TLR9 are 5'-CAGCTCCAGGGCCTATCGATTTTTGCACAGGACC-3' (SEQ ID NO:29) and 5'-GGTCCTGTGCAAAAATCGATAGGCCCTGGAGCTG-3' (SEQ ID NO:30). For

creating an expression vector containing the extracellular portion of human TLR3 connected to the TIR domain of human TLR9, the human TLR3 expression vector is cut with ClaI and limiting amounts of EcoRI and the fragment coding for the TIR domain of human TLR9 generated by a ClaI and EcoRI digestion of human TLR9 expression vector is ligated in the vector fragment containing the extracellular portion of hTLR3. Transfection into *E.coli* yields the expression vector hTLR3-TIR9 (human extracellular TLR3-human TLR9 TIR domain). The expressed product of hTLR3-TIR9 can interact with TLR3 ligands and also signal through an MyD88-mediated signal transduction pathway.

**hTLR9-TIR3**. A fusion construct with the extracellular domain of hTLR9 and the TIR domain of hTLR3 is prepared using an analogous strategy. For creating an expression vector containing the extracellular portion of human TLR9 connected to the TIR domain of human TLR3, the human TLR9 expression vector is cut with ClaI and limiting amounts of EcoRI and the fragment coding for the TIR domain of human TLR3 generated by a ClaI and EcoRI digestion of human TLR3 expression vector is ligated in the vector fragment containing the extracellular portion of hTLR9. Transfection into *E.coli* yields the expression vector hTLR9-TIR3 (human extracellular TLR9-human TLR3 TIR domain). The expressed product of hTLR9-TIR3 can interact with TLR9 ligands, e.g., CpG DNA, and signal through a signal transduction pathway in a manner like TLR3.

**Example 14. Sensitive in vitro Assay for Detecting Ligand Affinity Differences for a TLR**

Human 293 fibroblast cells stably transfected with murine TLR9 and an NF-κB-luciferase reporter were stimulated for 16 hours with the following fully phosphorothioated oligodeoxynucleotides (ODN):

| | | |
|---|---|---|
| 5890: | T*C*C*A*T*G*A*C*G*T*T*T*T*T*T*G*A*T*G*T*T | (SEQ ID NO:31) |
| 5895: | T*C*C*A*T*G*A*C*G*T*T*T*T*T*T*G*A*T*G | (SEQ ID NO:32) |
| 5896: | T*C*C*A*T*G*A*C*G*T*T*T*T*T*T*G*A | (SEQ ID NO:33) |
| 5897: | T*C*C*A*T*G*A*C*G*T*T*T*T*T*T | (SEQ ID NO:34) |

Concentration of the stimulus was titrated between 10 μM and 2 nM. The data is plotted in **Figure 6** as fold induction of NF-κB luciferase, relative to unstimulated

background, versus ODN concentration. The data displays typical first-order binding from which EC50 or maximal activity can be determined. EC50 is defined as the concentration of the ligand stimulus that results in 50% maximal activation. As shown in the figure, the EC50 ranges from 42 nM for ODN 5890 to 1220 nM for ODN 5897. The assay demonstrates sensitive differentiation between subtle changes in ligand.

**Example 15. Influence of Assay Kinetics on TLR Screening Assays**

Curves were prepared as in the previous Example 14 with the following ODN ligands, where * indicates phosphrothioate and _ indicates phosophodiester linkage:

| | | |
|---|---|---|
| 5890: | T*C*C*A*T*G*A*C*G*T*T*T*T*T*G*A*T*G*T*T | (SEQ ID NO:35) |
| 5497: | T*C*G*T*C*G*T*T*T*T_G_T_C_G_T*T*T*T*G*T*C*G*T*T | (SEQ ID NO:36) |
| 5746: | T*C_G*T*C_G*T*T*T*T_G*T*C_G*T*T*T*T*G*T*C_G*T*T | (SEQ ID NO:37) |
| 2006: | T*C*G*T*C*G*T*T*T*T*G*T*C*G*T*T*T*T*G*T*C*G*T*T | (SEQ ID NO:15) |
| 5902: | T*C*C*A*T*G*A*C_G_T*T*T*T*T*G*A*T_G*T*T | (SEQ ID NO:38) |

A family of stimulation curves was determined at various times of assay incubation between 1 and 24 hours. The EC50 was determined for each ligand at each time point. The EC50 was then plotted versus time to yield the resultant curves shown in **Figure 7**.

As evident from **Figure 7**, it is demonstrated that the kinetics of activation vary dependent on the ligand tested. Because luciferase has a three-hour half-life, the signal is transient and requires constant promoter-driven activation to be maintained. The maintenance is directly related to the signal delivered by the ligand/receptor complex. Thus analysis of time kinetics in such a fashion allows one to determine both affinity of ligand/receptor interaction and the availability of the ligand to the receptor through time. The principle is demonstrated as follows. The ODN 5890 is of higher affinity compared to the ODN 2006. When the ligand is made more labile to destruction by incorporating less stable diester linkages, the activity curves turn upward with time such as for ODN 5746, 5902 and 5497.

In the context of a screening assay for TLR/ligand interactions, limiting the assay to one time point would bias the assay. At 24 hours it would appear that only ODN 2006 and 5890 were ligand candidates, however this is clearly not the case. The assay also demonstrates that earlier time points, such as 6 hours in this example, would be the optimal time point for determining the greatest difference between

receptor/ligand affinities. Thus optimization of the screening assay can be adjusted depending on the desired information to be obtained from the screen, e.g., higher affinity of interaction versus stability and duration of receptor/ligand interaction.

**Figure 8** demonstrates the same principles shown with a murine TLR as in this example can be applied independent of the TLR utilized. For this set of data a 293 cell stably transfected with human TLR9 and NF-κB-luciferase was used.

**Example 16. Influence of Assay Kinetics on Maximal Activities in TLR Screening Assays**

Data was collected as in the previous Example 15, however the maximal activity (maximal fold induction) was plotted versus time in **Figures 9** and **10**. Such data analysis results in a prediction of biological efficacy. As can be seen from these figures, the lower affinity ODN, e.g., ODN 2006 and 5890 as demonstrated by the EC50 curves of Example 15, are clearly less efficient at delivering high activity.

**Example 17. Differential Outcomes of TLR Screening Assays Dependent on Promoter Utilization**

Human 293 fibroblast cells were transiently transfected with expression vector for TLR 7, TLR8, or TLR9 and one of the following reporter constructs bearing the following promoters driving the luciferase gene: NF-κB-luc, IP-10-luc, RANTES-luc, ISRE-luc, and IL-8-luc. The cells were stimulated for 16h with the maximal activity concentration of specific ligand. TLR9 was stimulated with CpG ODN 2006; TLR8 and TLR7 were stimulated with the imidazolquinalone R848. Results are shown in **Figure 11**. As evident from the figure, the promoter used influences the outcome of the screening assay dependent on the TLR in question. For example, NF-κB is a reliable marker for all TLRs tested, whereas in this set of experiments ISRE was only functional to some extent for TLR8. The IL-8 promoter is particularly sensitive for TLR7 or TLR8 screening assays but would be much less efficient in TLR9 assays.

What is claimed is:

## Claims

1. A screening method for identifying an immunostimulatory compound, comprising:

contacting a functional TLR3 with a test compound under conditions which, in absence of the test compound, permit a negative control response mediated by a TLR3 signal transduction pathway;

detecting a test response mediated by the TLR3 signal transduction pathway; and

determining the test compound is an immunostimulatory compound when the test response exceeds the negative control response.

2. A screening method for identifying an immunostimulatory compound, comprising:

contacting a functional TLR3 with a test compound under conditions which, in presence of a reference immunostimulatory compound, permit a reference response mediated by a TLR3 signal transduction pathway;

detecting a test response mediated by the TLR3 signal transduction pathway; and

determining the test compound is an immunostimulatory compound when the test response equals or exceeds the reference response.

3. A screening method for identifying a compound that modulates TLR3 signaling activity, comprising:

contacting a functional TLR3 with a test compound and a reference immunostimulatory compound under conditions which, in presence of the reference immunostimulatory compound alone, permit a reference response mediated by a TLR3 signal transduction pathway;

detecting a test-reference response mediated by the TLR3 signal transduction pathway;

determining the test compound is an agonist of TLR3 signaling activity when the test-reference response exceeds the reference response; and

determining the test compound is an antagonist of TLR3 signaling activity when the reference response exceeds the test-reference response.

4. A screening method for identifying species specificity of an immunostimulatory compound, comprising:

measuring a first species-specific response mediated by a TLR3 signal transduction pathway when a functional TLR3 of a first species is contacted with a test compound;

measuring a second species-specific response mediated by the TLR3 signal transduction pathway when a functional TLR3 of a second species is contacted with the test compound; and

comparing the first species-specific response with the second species-specific response.

5. The method of any one of claims 1-4, wherein the screening method is performed on a plurality of test compounds.

6. The method of claim 5, wherein the response mediated by the TLR3 signal transduction pathway is measured quantitatively.

7. The method of any one of claims 1-4, wherein the functional TLR3 is expressed in a cell.

8. The method of claim 7, wherein the cell is an isolated mammalian cell that naturally expresses the functional TLR3.

9. The method of claim 7, wherein the cell is an isolated mammalian cell that does not naturally express the functional TLR3, and wherein the cell comprises an expression vector for TLR3.

10. The method of claim 9, wherein the cell is a 293 human fibroblast.

11. The method of claim 7, wherein the cell comprises an expression vector comprising an isolated nucleic acid which encodes a reporter construct selected from the group of interleukin-6-luciferase (IL-6-luc), IL-8-luc, IL-12 p40-luc, IL-12 p40-β-Gal, NF-κB-luc, AP1-luc, IFN-α-luc, IFN-β-luc, RANTES-luc, TNF-luc, IP-10-luc, I-TAC-luc, and ISRE-luc.

12. The method of claim 11, wherein the reporter construct is ISRE-luc.

13. The method of any one of claims 1-4, wherein the functional TLR3 is part of a cell-free system.

14. The method of any one of claims 1-4, wherein the functional TLR3 is part of a complex with a non-TLR protein selected from the group consisting of MyD88, IL-1 receptor associated kinase 1-3 (IRAK1, IRAK2, IRAK3), tumor necrosis factor receptor-associated factor 1-6 (TRAF1 - TRAF6), IκB, NF-κB, MyD88-adapter-like (Mal), Toll-interleukin 1 receptor (TIR) domain-containing adapter protein (TIRAP), Tollip, Rac, and functional homologues and derivatives thereof.

15. The method of claim 14, wherein the non-TLR protein excludes MyD88.

16. The method of claim 2 or 3, wherein the reference immunostimulatory compound is a nucleic acid.

17. The method of claim 16, wherein the nucleic acid is a CpG nucleic acid.

18. The method of claim 2 or 3, wherein the reference immunostimulatory compound is a small molecule.

19. The method of any one of claims 1-4, wherein the test compound is a part of a combinatorial library of compounds.

20. The method of any one of claims 1-4, wherein the test compound is a nucleic acid.

21. The method of claim 20, wherein the nucleic acid is a CpG nucleic acid.

22. The method of any one of claims 1-4, wherein the test compound is a small molecule.

23. The method of any one of claims 1-4, wherein the test compound is a polypeptide.

24. The method of any one of claims 1-4, wherein the response mediated by a TLR3 signal transduction pathway is induction of a reporter gene under control of a promoter response element selected from the group consisting of ISRE, IL-6, IL-8, IL-12 p40, IFN-α, IFN-β, IFN-ω, RANTES, TNF, IP-10, and I-TAC.

25. The method of claim 24, wherein the reporter gene under control of a promoter response element is selected from the group consisting of ISRE-luc, IL-6-luc, IL-8-luc, IL-12 p40-luc, IL-12 p40-β-Gal, IFN-α-luc, IFN-β-luc, RANTES-luc, TNF-luc, IP-10-luc, and I-TAC-luc.

26. The method of claim 25, wherein the reporter gene under control of a promoter response element is ISRE-luc.

27. The method of claim 24, wherein the reporter gene is selected from the group consisting of IFN-α1-luc and IFN-α4-luc.

28. The method of any one of claims 1-4, wherein the response mediated by a TLR3 signal transduction pathway is selected from the group consisting of (a) induction of a reporter gene under control of a minimal promoter responsive to a transcription factor selected from the group consisting of AP1, NF-κB, ATF2, IRF3, and IRF7; (b) secretion of a chemokine; and (c) secretion of a cytokine.

29. The method of claim 28, wherein the response mediated by a TLR3 signal transduction pathway is induction of a reporter gene selected from the group consisting of AP1-luc and NF-κB-luc.

30. The method of claim 28, wherein the response mediated by a TLR3 signal transduction pathway is secretion of a type 1 IFN.

31. The method of claim 28, wherein the response mediated by a TLR3 signal transduction pathway is secretion of a chemokine selected from the group consisting of CCL5 (RANTES), CXCL9 (Mig), CXCL10 (IP-10), and CXCL11 (I-TAC).

32. The method of any one of claims 1-3, wherein the contacting a functional TLR3 with a test compound further comprises, for each test compound, contacting with the test compound at each of a plurality of concentrations.

33. The method of any one of claims 1-3, wherein the detecting is performed 6-12 hours following the contacting.

34. The method of any one of claims 1-3, wherein the detecting is performed 16-24 hours following the contacting.

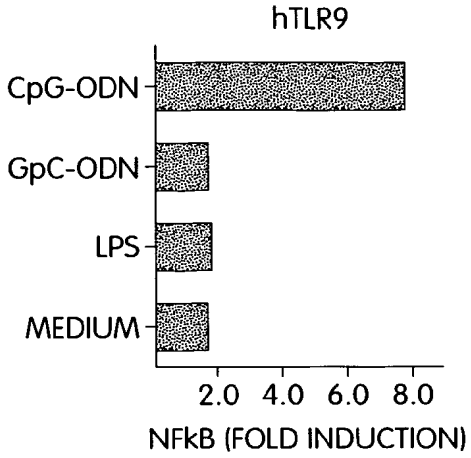

hTLR9
CpG-ODN
GpC-ODN
LPS
MEDIUM
2.0 4.0 6.0 8.0
NFkB (FOLD INDUCTION)

Fig. 1A

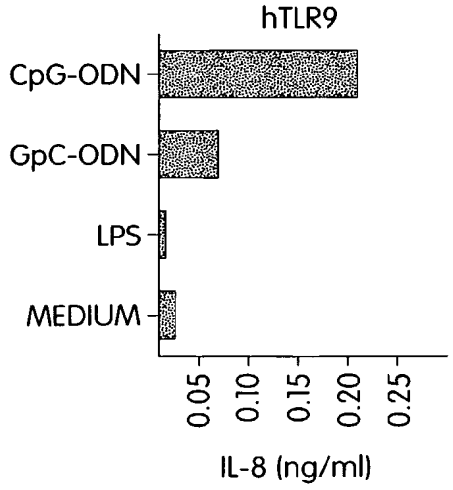

hTLR9
CpG-ODN
GpC-ODN
LPS
MEDIUM
0.05 0.10 0.15 0.20 0.25
IL-8 (ng/ml)

Fig. 1B

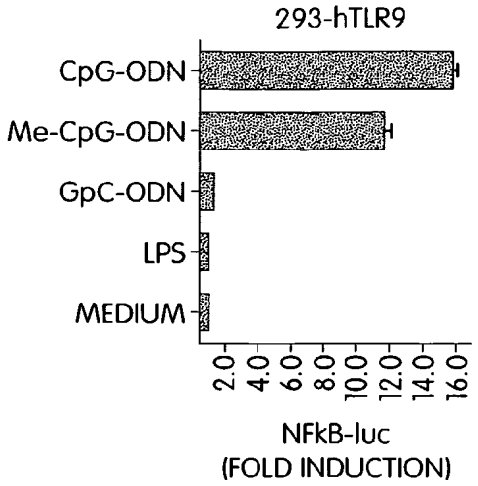



Fig. 2

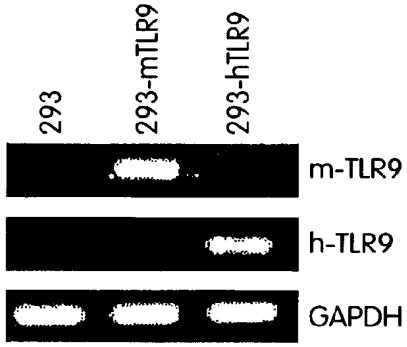



Fig. 3

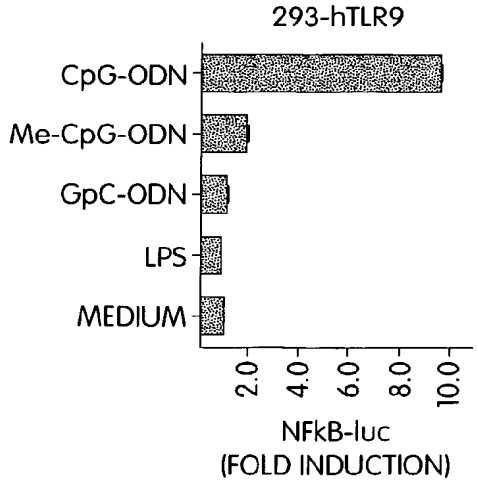



Fig. 4

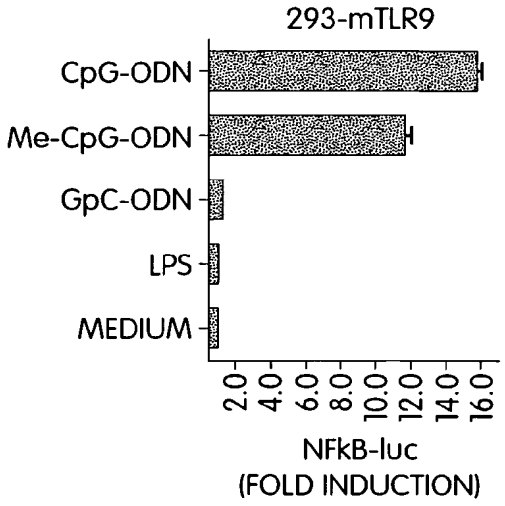

293-mTLR9
CpG-ODN
Me-CpG-ODN
GpC-ODN
LPS
MEDIUM
2.0
4.0
6.0
8.0
10.0
12.0
14.0
16.0
NFkB-luc
(FOLD INDUCTION)

Fig. 5

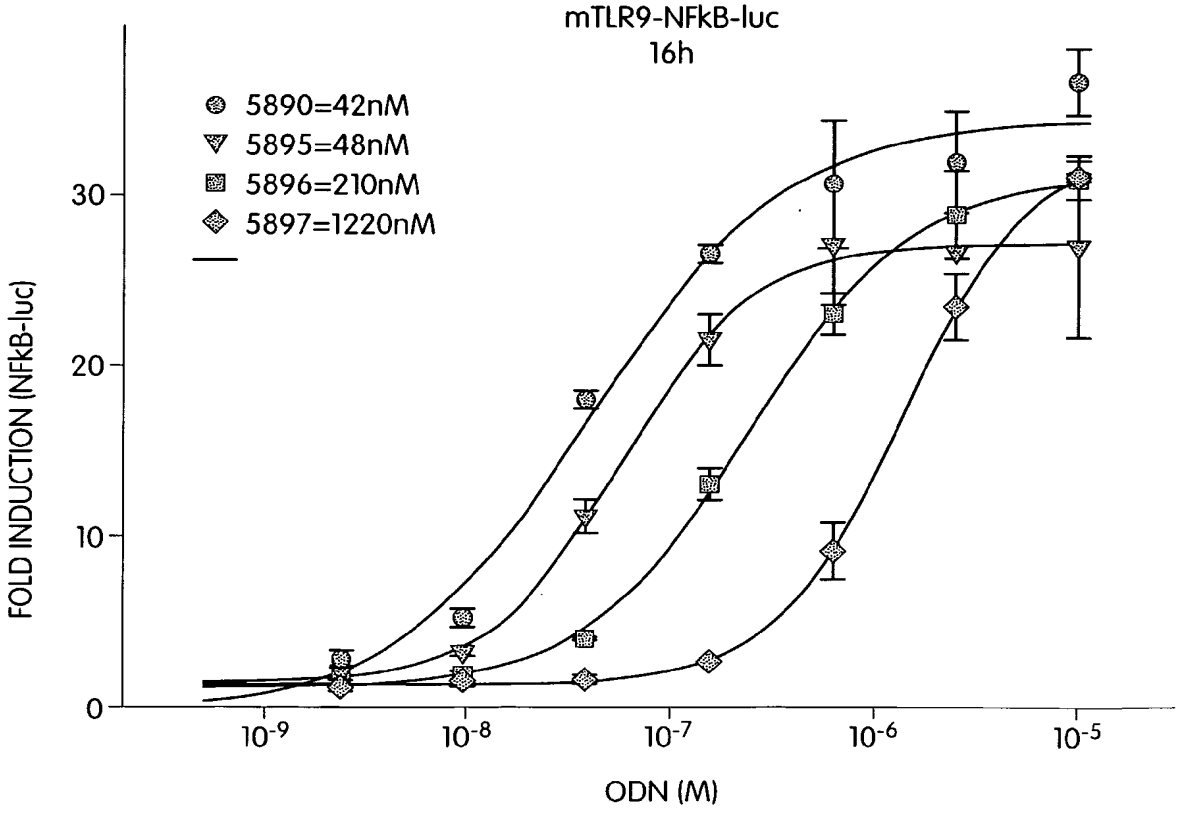

mTLR9-NFkB-luc
16h
5890=42nM
5895=48nM
5896=210nM
5897=1220nM
FOLD INDUCTION (NFkB-luc)
0
10
20
30
10-9
10-8
10-7
10-6
10-5
ODN (M)

Fig. 6

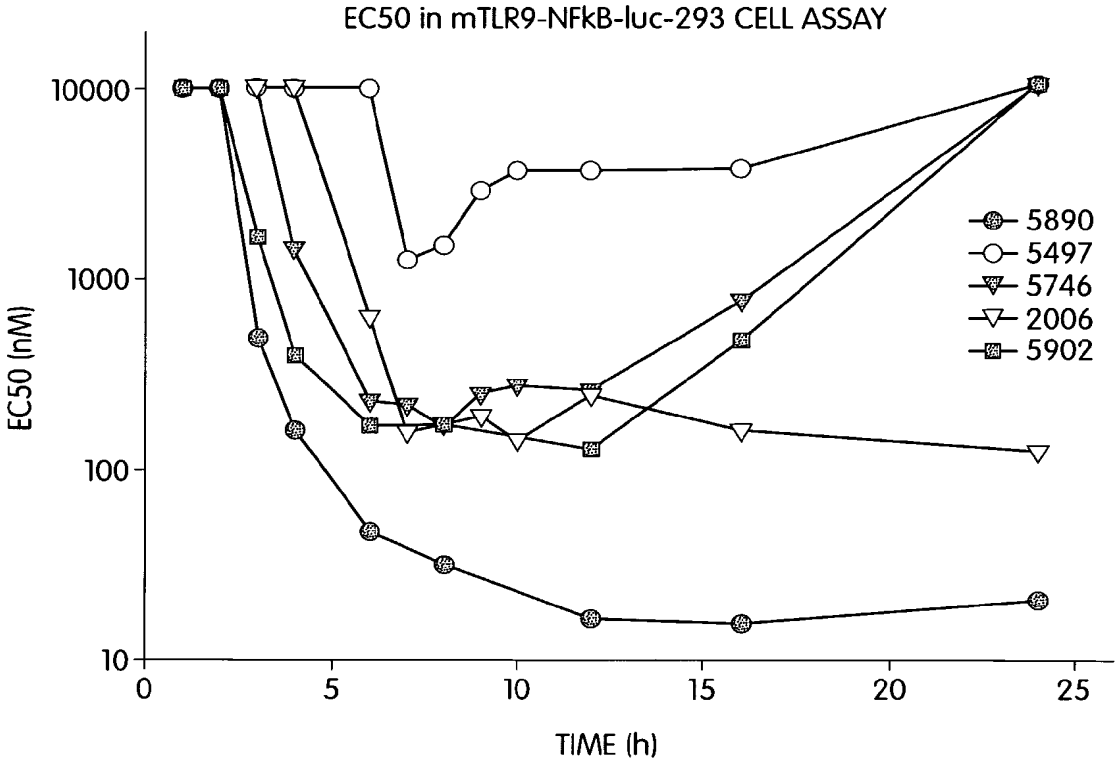

EC50 in mTLR9-NFkB-luc-293 CELL ASSAY
EC50 (nM)
10000
1000
100
10
0
5
10
15
20
25
TIME (h)
5890
5497
5746
2006
5902

Fig. 7

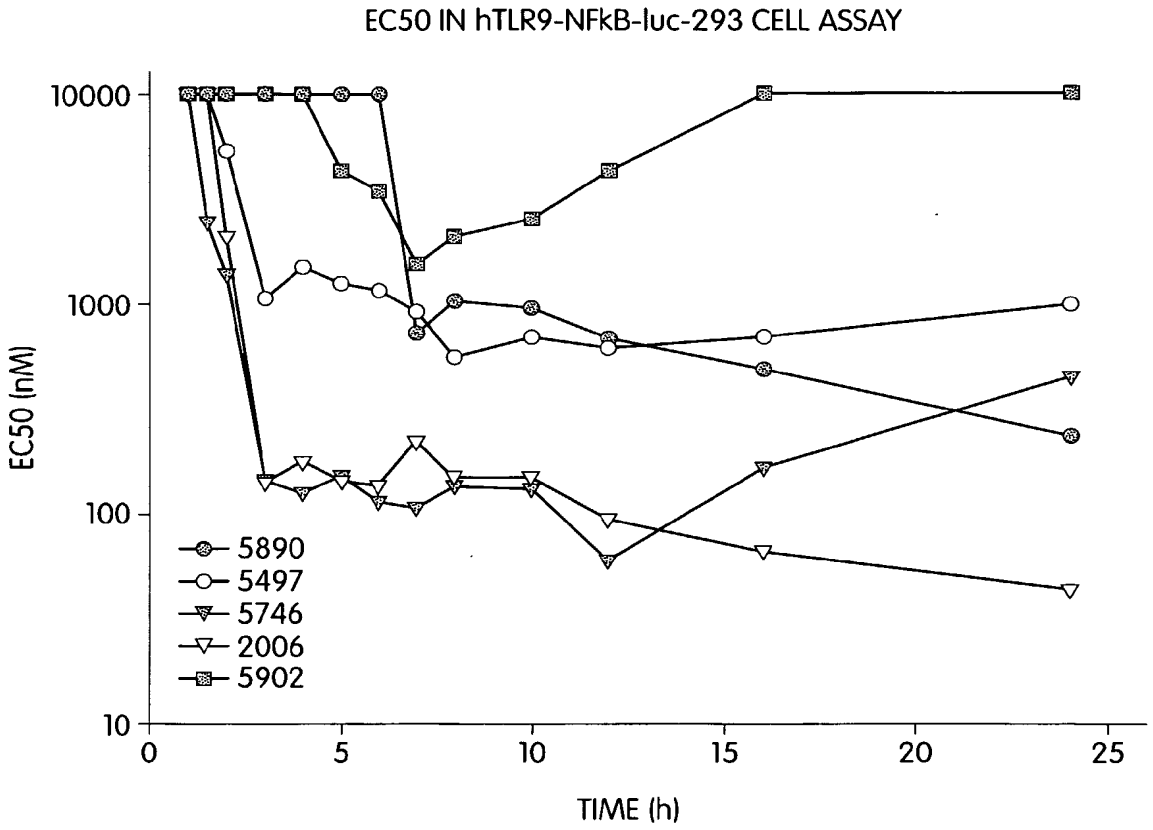

EC50 IN hTLR9-NFkB-luc-293 CELL ASSAY
10000
1000
100
10
EC50 (nM)
0
5
10
15
20
25
TIME (h)
5890
5497
5746
2006
5902

Fig. 8

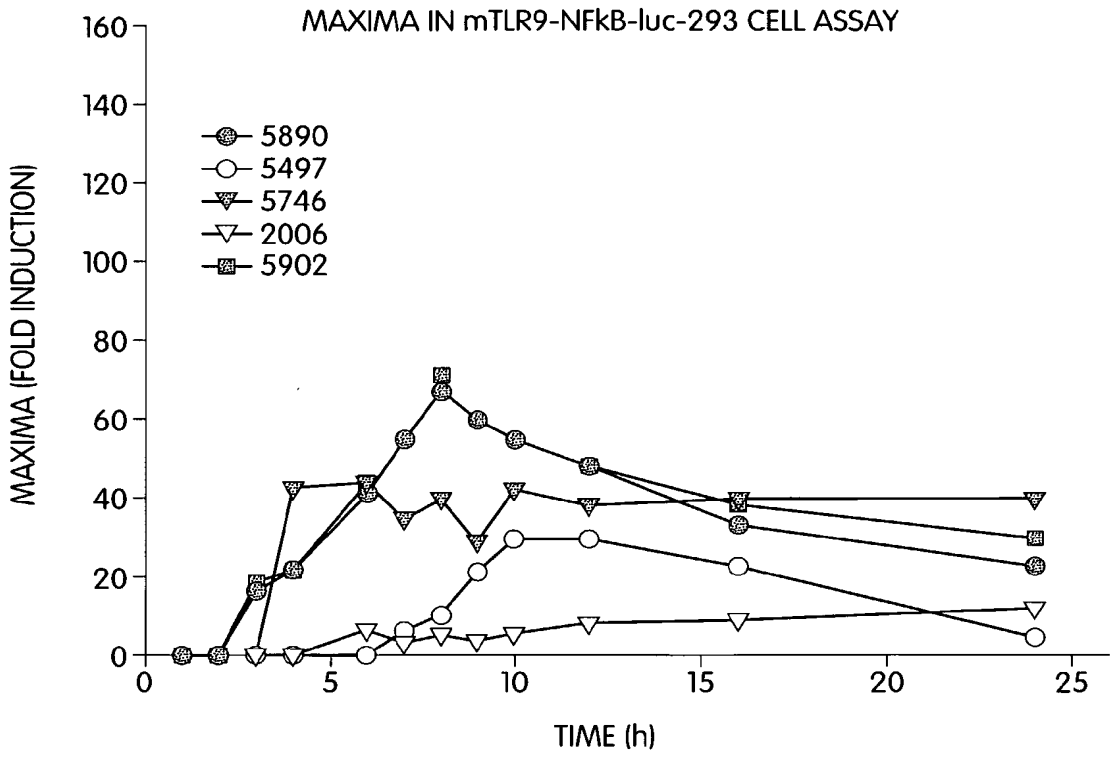

MAXIMA IN mTLR9-NFkB-luc-293 CELL ASSAY
MAXIMA (FOLD INDUCTION)
TIME (h)
5890
5497
5746
2006
5902
0
20
40
60
80
100
120
140
160
0
5
10
15
20
25

Fig. 9

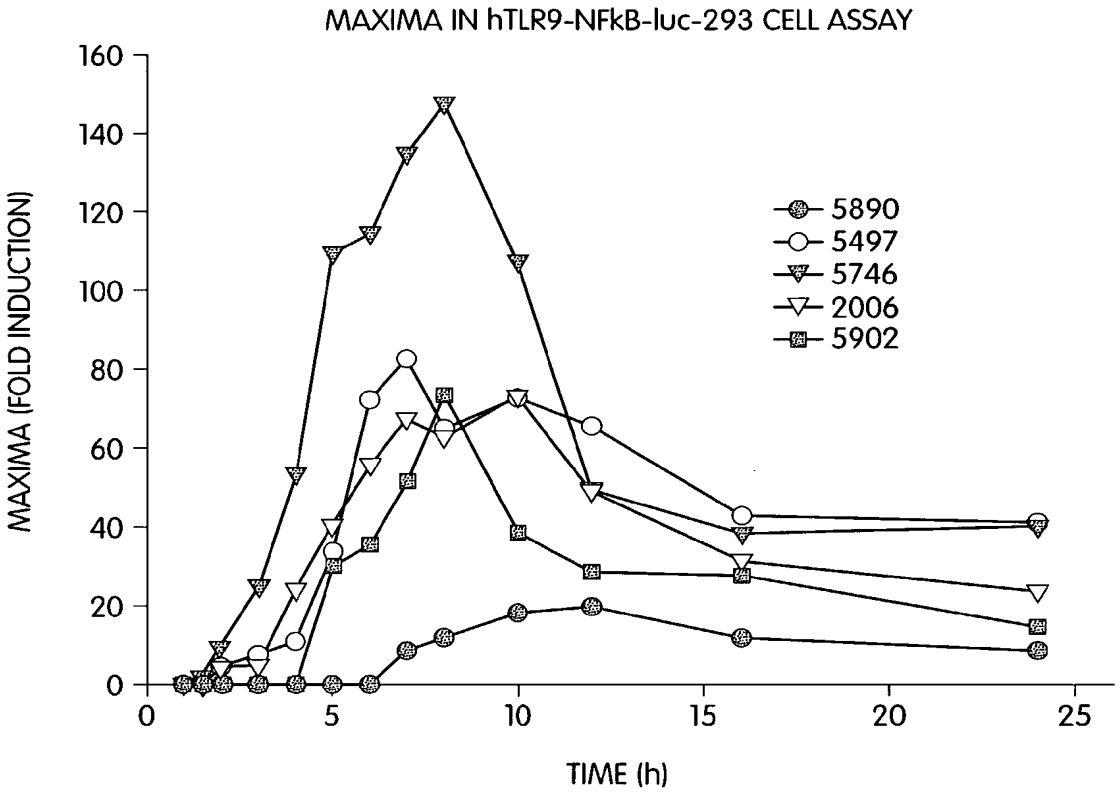

MAXIMA IN hTLR9-NFkB-luc-293 CELL ASSAY
MAXIMA (FOLD INDUCTION)
160
140
120
100
80
60
40
20
0
0
5
10
15
20
25
TIME (h)
5890
5497
5746
2006
5902

Fig. 10

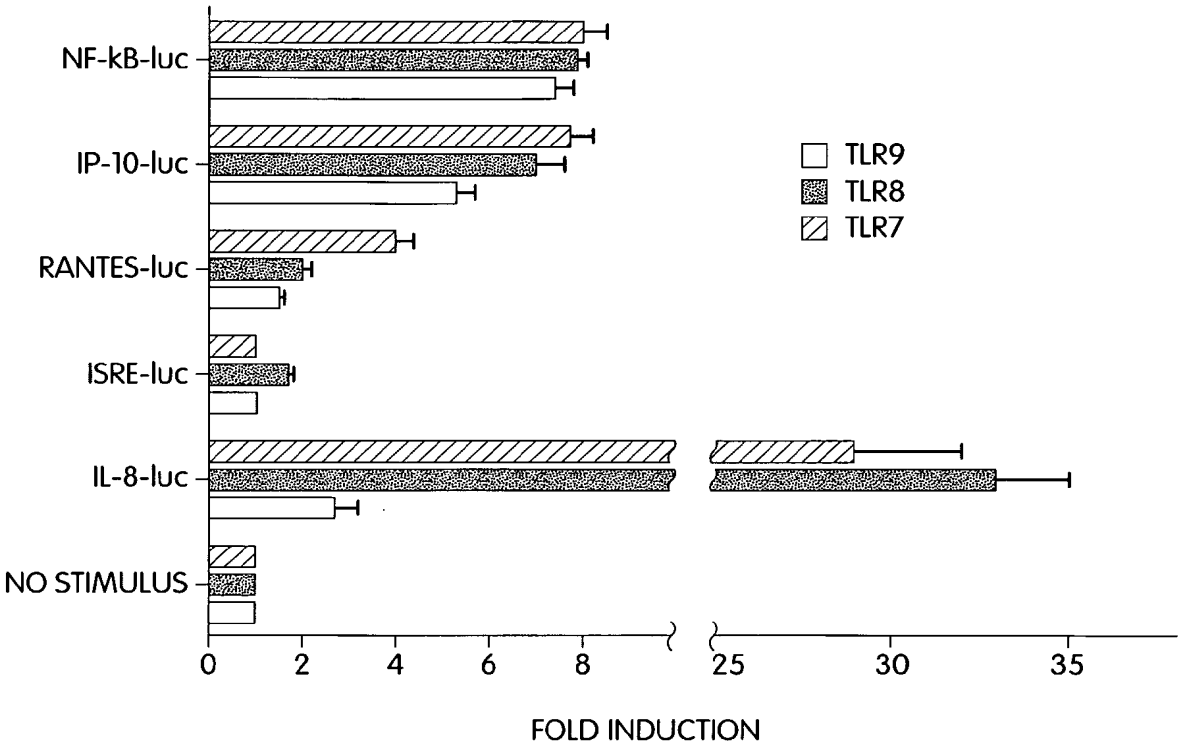

NF-kB-luc
IP-10-luc
RANTES-luc
ISRE-luc
IL-8-luc
NO STIMULUS
TLR9
TLR8
TLR7
0
2
4
6
8
25
30
35
FOLD INDUCTION

Fig. 11

SEQUENCE LISTING

<110> Coley Pharmaceutical GmbH

<120> TOLL-LIKE RECEPTOR 3 SIGNALING AGONISTS AND ANTAGONISTS

<130> C01041.70031

<160> 117

<170> PatentIn version 3.1

<210> 1
<211> 3029
<212> DNA
<213> Homo sapiens

<400> 1
gcggccgcgt cgacgaaatg tctggatttg gactaaagaa aaaaggaaag gctagcagtc 60

atccaacaga atcatgagac agactttgcc ttgtatctac ttttgggggg gccttttgcc 120

ctttgggatg ctgtgtgcat cctccaccac caagtgcact gttagccatg aagttgctga 180

ctgcagccac ctgaagttga ctcaggtacc cgatgatcta cccacaaaca taacagtgtt 240

gaaccttacc cataatcaac tcagaagatt accagccgcc aacttcacaa ggtatagcca 300

gctaactagc ttggatgtag gatttaacac catctcaaaa ctggagccag aattgtgcca 360

gaaacttccc atgttaaaag ttttgaacct ccagcacaat gagctatctc aactttctga 420

taaaaccttt gccttctgca cgaatttgac tgaactccat ctcatgtcca actcaatcca 480

gaaaattaaa aataatccct ttgtcaagca gaagaattta atcacattag atctgtctca 540

taatggcttg tcatctacaa aattaggaac tcaggttcag ctggaaaatc tccaagagct 600

tctattatca aacaataaaa ttcaagcgct aaaaagtgaa gaactggata tctttgccaa 660

ttcatcttta aaaaaattag agttgtcatc gaatcaaatt aaagagtttt ctccagggtg 720

ttttcacgca attggaagat tatttggcct ctttctgaac aatgtccagc tgggtcccag 780

ccttacagag aagctatgtt tggaattagc aaacacaagc attcggaatc tgtctctgag 840

taacagccag ctgtccacca ccagcaatac aactttcttg ggactaaagt ggacaaatct 900

cactatgctc gatctttcct acaacaactt aaatgtggtt ggtaacgatt cctttgcttg 960

gcttccacaa ctagaatatt tcttcctaga gtataataat atacagcatt tgttttctca 1020

ctctttgcac gggcttttca atgtgaggta cctgaatttg aaacggtctt ttactaaaca 1080

aagtatttcc cttgcctcac tccccaagat tgatgatttt tcttttcagt ggctaaaatg 1140

tttggagcac cttaacatgg aagataatga tattccaggc ataaaaagca atatgttcac 1200

aggattgata aacctgaaat acttaagtct atccaactcc tttacaagtt tgcgaacttt 1260

gacaaatgaa acatttgtat cacttgctca ttctccctta cacatactca acctaaccaa 1320

gaataaaatc tcaaaaatag agagtgatgc tttctcttgg ttgggccacc tagaagtact 1380

tgacctgggc cttaatgaaa ttgggcaaga actcacaggc caggaatgga gaggtctaga 1440

aaatattttc gaaatctatc tttcctacaa caagtacctg cagctgacta ggaactcctt 1500

tgccttggtc ccaagccttc aacgactgat gctccgaagg gtggccctta aaaatgtgga 1560

tagctctcct tcaccattcc agcctcttcg taacttgacc attctggatc taagcaacaa 1620

caacatagcc aacataaatg atgacatgtt ggagggtctt gagaaactag aaattctcga 1680

tttgcagcat aacaacttag cacggctctg gaaacacgca aaccctggtg gtcccattta 1740

tttcctaaag ggtctgtctc acctccacat ccttaacttg gagtccaacg gctttgacga 1800

gatcccagtt gaggtcttca aggatttatt tgaactaaag atcatcgatt taggattgaa 1860

taatttaaac acacttccag catctgtctt taataatcag gtgtctctaa agtcattgaa 1920

ccttcagaag aatctcataa catccgttga gaagaaggtt ttcgggccag ctttcaggaa 1980

cctgactgag ttagatatgc gctttaatcc ctttgattgc acgtgtgaaa gtattgcctg 2040

gtttgttaat tggattaacg agacccatac caacatccct gagctgtcaa gccactacct 2100

ttgcaacact ccacctcact atcatgggtt cccagtgaga ctttttgata catcatcttg 2160

caaagacagt gccccctttg aactcttttt catgatcaat accagtatcc tgttgatttt 2220

tatctttatt gtacttctca tccactttga gggctggagg atatcttttt attggaatgt 2280

ttcagtacat cgagttcttg gtttcaaaga aatagacaga cagacagaac agtttgaata 2340

tgcagcatat ataattcatg cctataaaga taaggattgg gtctgggaac atttctcttc 2400

aatggaaaag gaagaccaat ctctcaaatt ttgtctggaa gaaagggact ttgaggcggg 2460

tgtttttgaa ctagaagcaa ttgttaacag catcaaaaga agcagaaaaa ttatttttgt 2520

tataacacac catctattaa aagacccatt atgcaaaaga ttcaaggtac atcatgcagt 2580

tcaacaagct attgaacaaa atctggattc cattatattg gttttccttg aggagattcc 2640

agattataaa ctgaaccatg cactctgttt gcgaagagga atgtttaaat ctcactgcat 2700

cttgaactgg ccagttcaga aagaacggat aggtgccttt cgtcataaat tgcaagtagc 2760

acttggatcc aaaaactctg tacattaaat ttatttaaat attcaattag caaaggagaa 2820

actttctcaa tttaaaaagt tctatggcaa atttaagttt tccataaagg tgttataatt 2880

tgtttattca tatttgtaaa tgattatatt ctatcacaat tacatctctt ctaggaaaat 2940

gtgtctcctt atttcaggcc tatttttgac aattgactta attttaccca aaataaaaca 3000

tataagcacg caaaaaaaaa aaaaaaaaa 3029

<210> 2
<211> 904
<212> PRT
<213> Homo sapiens

<400> 2

Met Arg Gln Thr Leu Pro Cys Ile Tyr Phe Trp Gly Gly Leu Leu Pro
1 5 10 15

Phe Gly Met Leu Cys Ala Ser Ser Thr Thr Lys Cys Thr Val Ser His
20 25 30

Glu Val Ala Asp Cys Ser His Leu Lys Leu Thr Gln Val Pro Asp Asp
35 40 45

Leu Pro Thr Asn Ile Thr Val Leu Asn Leu Thr His Asn Gln Leu Arg
50 55 60

Arg Leu Pro Ala Ala Asn Phe Thr Arg Tyr Ser Gln Leu Thr Ser Leu
65 70 75 80

Asp Val Gly Phe Asn Thr Ile Ser Lys Leu Glu Pro Glu Leu Cys Gln
85 90 95

Lys Leu Pro Met Leu Lys Val Leu Asn Leu Gln His Asn Glu Leu Ser
100 105 110

Gln Leu Ser Asp Lys Thr Phe Ala Phe Cys Thr Asn Leu Thr Glu Leu
115 120 125

His Leu Met Ser Asn Ser Ile Gln Lys Ile Lys Asn Asn Pro Phe Val
130 135 140

Lys Gln Lys Asn Leu Ile Thr Leu Asp Leu Ser His Asn Gly Leu Ser
145 150 155 160

Ser Thr Lys Leu Gly Thr Gln Val Gln Leu Glu Asn Leu Gln Glu Leu
165 170 175

Leu Leu Ser Asn Asn Lys Ile Gln Ala Leu Lys Ser Glu Glu Leu Asp
180 185 190

Ile Phe Ala Asn Ser Ser Leu Lys Lys Leu Glu Leu Ser Ser Asn Gln

195 200 205

Ile Lys Glu Phe Ser Pro Gly Cys Phe His Ala Ile Gly Arg Leu Phe
210 215 220

Gly Leu Phe Leu Asn Asn Val Gln Leu Gly Pro Ser Leu Thr Glu Lys
225 230 235 240

Leu Cys Leu Glu Leu Ala Asn Thr Ser Ile Arg Asn Leu Ser Leu Ser
245 250 255

Asn Ser Gln Leu Ser Thr Thr Ser Asn Thr Thr Phe Leu Gly Leu Lys
260 265 270

Trp Thr Asn Leu Thr Met Leu Asp Leu Ser Tyr Asn Asn Leu Asn Val
275 280 285

Val Gly Asn Asp Ser Phe Ala Trp Leu Pro Gln Leu Glu Tyr Phe Phe
290 295 300

Leu Glu Tyr Asn Asn Ile Gln His Leu Phe Ser His Ser Leu His Gly
305 310 315 320

Leu Phe Asn Val Arg Tyr Leu Asn Leu Lys Arg Ser Phe Thr Lys Gln
325 330 335

Ser Ile Ser Leu Ala Ser Leu Pro Lys Ile Asp Asp Phe Ser Phe Gln
340 345 350

Trp Leu Lys Cys Leu Glu His Leu Asn Met Glu Asp Asn Asp Ile Pro
355 360 365

Gly Ile Lys Ser Asn Met Phe Thr Gly Leu Ile Asn Leu Lys Tyr Leu
370 375 380

Ser Leu Ser Asn Ser Phe Thr Ser Leu Arg Thr Leu Thr Asn Glu Thr
385 390 395 400

Phe Val Ser Leu Ala His Ser Pro Leu His Ile Leu Asn Leu Thr Lys
405 410 415

Asn Lys Ile Ser Lys Ile Glu Ser Asp Ala Phe Ser Trp Leu Gly His
420 425 430

```
Leu Glu Val Leu Asp Leu Gly Leu Asn Glu Ile Gly Gln Glu Leu Thr
        435                 440                 445

Gly Gln Glu Trp Arg Gly Leu Glu Asn Ile Phe Glu Ile Tyr Leu Ser
    450                 455                 460

Tyr Asn Lys Tyr Leu Gln Leu Thr Arg Asn Ser Phe Ala Leu Val Pro
465                 470                 475                 480

Ser Leu Gln Arg Leu Met Leu Arg Arg Val Ala Leu Lys Asn Val Asp
                485                 490                 495

Ser Ser Pro Ser Pro Phe Gln Pro Leu Arg Asn Leu Thr Ile Leu Asp
            500                 505                 510

Leu Ser Asn Asn Asn Ile Ala Asn Ile Asn Asp Asp Met Leu Glu Gly
        515                 520                 525

Leu Glu Lys Leu Glu Ile Leu Asp Leu Gln His Asn Asn Leu Ala Arg
    530                 535                 540

Leu Trp Lys His Ala Asn Pro Gly Gly Pro Ile Tyr Phe Leu Lys Gly
545                 550                 555                 560

Leu Ser His Leu His Ile Leu Asn Leu Glu Ser Asn Gly Phe Asp Glu
                565                 570                 575

Ile Pro Val Glu Val Phe Lys Asp Leu Phe Glu Leu Lys Ile Ile Asp
            580                 585                 590

Leu Gly Leu Asn Asn Leu Asn Thr Leu Pro Ala Ser Val Phe Asn Asn
        595                 600                 605

Gln Val Ser Leu Lys Ser Leu Asn Leu Gln Lys Asn Leu Ile Thr Ser
    610                 615                 620

Val Glu Lys Lys Val Phe Gly Pro Ala Phe Arg Asn Leu Thr Glu Leu
625                 630                 635                 640

Asp Met Arg Phe Asn Pro Phe Asp Cys Thr Cys Glu Ser Ile Ala Trp
                645                 650                 655

Phe Val Asn Trp Ile Asn Glu Thr His Thr Asn Ile Pro Glu Leu Ser
            660                 665                 670
```

```
Ser His Tyr Leu Cys Asn Thr Pro Pro His Tyr His Gly Phe Pro Val
        675                 680                 685

Arg Leu Phe Asp Thr Ser Ser Cys Lys Asp Ser Ala Pro Phe Glu Leu
    690                 695                 700

Phe Phe Met Ile Asn Thr Ser Ile Leu Leu Ile Phe Ile Phe Ile Val
705                 710                 715                 720

Leu Leu Ile His Phe Glu Gly Trp Arg Ile Ser Phe Tyr Trp Asn Val
                725                 730                 735

Ser Val His Arg Val Leu Gly Phe Lys Glu Ile Asp Arg Gln Thr Glu
            740                 745                 750

Gln Phe Glu Tyr Ala Ala Tyr Ile Ile His Ala Tyr Lys Asp Lys Asp
        755                 760                 765

Trp Val Trp Glu His Phe Ser Ser Met Glu Lys Glu Asp Gln Ser Leu
    770                 775                 780

Lys Phe Cys Leu Glu Glu Arg Asp Phe Glu Ala Gly Val Phe Glu Leu
785                 790                 795                 800

Glu Ala Ile Val Asn Ser Ile Lys Arg Ser Arg Lys Ile Ile Phe Val
                805                 810                 815

Ile Thr His His Leu Leu Lys Asp Pro Leu Cys Lys Arg Phe Lys Val
            820                 825                 830

His His Ala Val Gln Gln Ala Ile Glu Gln Asn Leu Asp Ser Ile Ile
        835                 840                 845

Leu Val Phe Leu Glu Glu Ile Pro Asp Tyr Lys Leu Asn His Ala Leu
    850                 855                 860

Cys Leu Arg Arg Gly Met Phe Lys Ser His Cys Ile Leu Asn Trp Pro
865                 870                 875                 880

Val Gln Lys Glu Arg Ile Gly Ala Phe Arg His Lys Leu Gln Val Ala
                885                 890                 895

Leu Gly Ser Lys Asn Ser Val His
            900
```

<210> 3
<211> 3310
<212> DNA
<213> Mus musculus

<400> 3
tagaatatga tacagggatt gcacccataa tctgggctga atcatgaaag ggtgttcctc 60

ttatctaatg tactcctttg ggggactttt gtccctatgg attcttctgg tgtcttccac 120

aaaccaatgc actgtgagat acaacgtagc tgactgcagc catttgaagc taacacacat 180

acctgatgat cttccctcta acataacagt gttgaatctt actcacaacc aactcagaag 240

attaccacct accaacttta caagatacag ccaacttgct atcttggatg caggatttaa 300

ctccatttca aaactggagc cagaactgtg ccaaatactc cctttgttga aagtattgaa 360

cctgcaacat aatgagctct ctcagatttc tgatcaaacc tttgtcttct gcacgaacct 420

gacagaactc gatctaatgt ctaactcaat acacaaaatt aaaagcaacc ctttcaaaaa 480

ccagaagaat ctaatcaaat tagatttgtc tcataatggt ttatcatcta caaagttggg 540

aacgggggtc caactggaga acctccaaga actgctctta gcaaaaaata aaatccttgc 600

gttgcgaagt gaagaacttg agtttcttgg caattcttct ttacgaaagt tggacttgtc 660

atcaaatcca cttaaagagt tctccccggg gtgtttccag acaattggca agttattcgc 720

cctcctcttg aacaacgccc aactgaaccc ccacctcaca gagaagcttt gctgggaact 780

ttcaaacaca agcatccaga atctctctct ggctaacaac cagctgctgg ccaccagcga 840

gagcactttc tctgggctga agtggacaaa tctcacccag ctcgatcttt cctacaacaa 900

cctccatgat gtcggcaacg gttccttctc ctatctccca agcctgaggt atctgtctct 960

ggagtacaac aatatacagc gtctgtcccc tcgctctttt tatggactct ccaacctgag 1020

gtacctgagt ttgaagcgag catttactaa gcaaagtgtt tcacttgctt cacatcccaa 1080

cattgacgat ttttcctttc aatggttaaa atatttggaa tatctcaaca tggatgacaa 1140

taatattcca agtaccaaaa gcaatacctt cacgggattg gtgagtctga agtacctaag 1200

tctttccaaa actttcacaa gtttgcaaac tttaacaaat gaaacatttg tgtcacttgc 1260

tcattctccc ttgctcactc tcaacttaac gaaaaatcac atctcaaaaa tagcaaatgg 1320

tactttctct tggttaggcc aactcaggat acttgatctc ggccttaatg aaattgaaca 1380

aaaactcagc ggccaggaat ggagaggtct gagaaatata tttgagatct acctatccta 1440

taacaaatac ctccaactgt ctaccagttc ctttgcattg gtccccagcc ttcaaagact 1500

gatgctcagg agggtggccc ttaaaaatgt ggatatctcc ccttcacctt tccgccctct 1560

tcgtaacttg accattctgg acttaagcaa caacaacata gccaacataa atgaggactt 1620

gctggagggt cttgagaatc tagaaatcct ggattttcag cacaataact tagccaggct 1680

ctggaaacgc gcaaaccccg gtggtcccgt taatttcctg aaggggctgt ctcacctcca 1740

catcttgaat ttagagtcca acggcttaga tgaaatccca gtcggggttt tcaagaactt 1800

attcgaacta aagagcatca atctaggact gaataactta aacaaacttg aaccattcat 1860

ttttgatgac cagacatctc taaggtcact gaacctccag aagaacctca taacatctgt 1920

tgagaaggat gttttcgggc cgccttttca aaacctgaac agtttagata tgcgcttcaa 1980

tccgttcgac tgcacgtgtg aaagtatttc ctggtttgtt aactggatca accagaccca 2040

cactaatatc tttgagctgt ccactcacta cctctgtaac actccacatc attattatgg 2100

cttccccctg aagcttttcg atacatcatc ctgtaaagac agcgcccccc ttgaactcct 2160

cttcataatc agcaccagta tgctcctggt ttttatactt gtggtactgc tcattcacat 2220

cgagggctgg aggatctctt tttactggaa tgtttcagtg catcggattc ttggtttcaa 2280

ggaaatagac acacaggctg agcagtttga atatacagcc tacataattc atgcccataa 2340

agacagagac tgggtctggg aacatttctc cccaatggaa gaacaagacc aatctctcaa 2400

attttgccta gaagaaaggg actttgaagc aggcgtcctt ggacttgaag caattgttaa 2460

tagcatcaaa agaagccgaa aaatcatttt cgttatcaca caccatttat taaaagaccc 2520

tctgtgcaga agattcaagg tacatcacgc agttcagcaa gctattgagc aaaatctgga 2580

ttcaattata ctgatttttc tccagaatat tccagattat aaactaaacc atgcactctg 2640

tttgcgaaga ggaatgttta aatctcattg catcttgaac tggccagttc agaaagaacg 2700

gataaatgcc tttcatcata aattgcaagt agcacttgga tctcggaatt cagcacatta 2760

aactcatttg aagatttgga gtcggtaaag ggatagatcc aatttataaa ggtccatcat 2820

gaatctaagt tttacttgaa agttttgtat atttatttat atgtatagat gatgatatta 2880

catcacaatc caatctcagt tttgaaatat ttcggcttat ttcattgaca tctggtttat 2940

tcactccaaa taaacacatg ggcagttaaa aacatcctct attaatagat tacccattaa 3000

ttcttgaggt gtatcacagc tttaaagggt tttaaatatt tttatataaa taagactgag 3060

agttttataa atgtaatttt ttaaaactcg agtcttactg tgtagctcag aaaggcctgg 3120

aaattaatat attagagagt catgtcttga acttatttat ctctgcctcc ctctgtctcc 3180

agagtgttgc ttttaagggc atgtagcacc acacccagct atgtacgtgt gggattttat 3240

aatgctcatt tttgagacgt ttatagaata aaagataatt gcttttatgg tataaggcta 3300

cttgaggtaa 3310

<210> 4
<211> 905
<212> PRT
<213> Mus musculus

<400> 4

Met Lys Gly Cys Ser Ser Tyr Leu Met Tyr Ser Phe Gly Gly Leu Leu
1 5 10 15

Ser Leu Trp Ile Leu Leu Val Ser Ser Thr Asn Gln Cys Thr Val Arg
20 25 30

Tyr Asn Val Ala Asp Cys Ser His Leu Lys Leu Thr His Ile Pro Asp
35 40 45

Asp Leu Pro Ser Asn Ile Thr Val Leu Asn Leu Thr His Asn Gln Leu
50 55 60

Arg Arg Leu Pro Pro Thr Asn Phe Thr Arg Tyr Ser Gln Leu Ala Ile
65 70 75 80

Leu Asp Ala Gly Phe Asn Ser Ile Ser Lys Leu Glu Pro Glu Leu Cys
85 90 95

Gln Ile Leu Pro Leu Leu Lys Val Leu Asn Leu Gln His Asn Glu Leu
100 105 110

Ser Gln Ile Ser Asp Gln Thr Phe Val Phe Cys Thr Asn Leu Thr Glu
115 120 125

Leu Asp Leu Met Ser Asn Ser Ile His Lys Ile Lys Ser Asn Pro Phe
130 135 140

Lys Asn Gln Lys Asn Leu Ile Lys Leu Asp Leu Ser His Asn Gly Leu
145 150 155 160

Ser Ser Thr Lys Leu Gly Thr Gly Val Gln Leu Glu Asn Leu Gln Glu
165 170 175

Leu Leu Leu Ala Lys Asn Lys Ile Leu Ala Leu Arg Ser Glu Glu Leu
180 185 190

Glu Phe Leu Gly Asn Ser Ser Leu Arg Lys Leu Asp Leu Ser Ser Asn

```
        195                 200                 205

Pro Leu Lys Glu Phe Ser Pro Gly Cys Phe Gln Thr Ile Gly Lys Leu
    210                 215                 220

Phe Ala Leu Leu Leu Asn Asn Ala Gln Leu Asn Pro His Leu Thr Glu
225                 230                 235                 240

Lys Leu Cys Trp Glu Leu Ser Asn Thr Ser Ile Gln Asn Leu Ser Leu
                245                 250                 255

Ala Asn Asn Gln Leu Leu Ala Thr Ser Glu Ser Thr Phe Ser Gly Leu
            260                 265                 270

Lys Trp Thr Asn Leu Thr Gln Leu Asp Leu Ser Tyr Asn Asn Leu His
        275                 280                 285

Asp Val Gly Asn Gly Ser Phe Ser Tyr Leu Pro Ser Leu Arg Tyr Leu
    290                 295                 300

Ser Leu Glu Tyr Asn Asn Ile Gln Arg Leu Ser Pro Arg Ser Phe Tyr
305                 310                 315                 320

Gly Leu Ser Asn Leu Arg Tyr Leu Ser Leu Lys Arg Ala Phe Thr Lys
                325                 330                 335

Gln Ser Val Ser Leu Ala Ser His Pro Asn Ile Asp Asp Phe Ser Phe
            340                 345                 350

Gln Trp Leu Lys Tyr Leu Glu Tyr Leu Asn Met Asp Asp Asn Asn Ile
        355                 360                 365

Pro Ser Thr Lys Ser Asn Thr Phe Thr Gly Leu Val Ser Leu Lys Tyr
    370                 375                 380

Leu Ser Leu Ser Lys Thr Phe Thr Ser Leu Gln Thr Leu Thr Asn Glu
385                 390                 395                 400

Thr Phe Val Ser Leu Ala His Ser Pro Leu Leu Thr Leu Asn Leu Thr
                405                 410                 415

Lys Asn His Ile Ser Lys Ile Ala Asn Gly Thr Phe Ser Trp Leu Gly
            420                 425                 430
```

```
Gln Leu Arg Ile Leu Asp Leu Gly Leu Asn Glu Ile Glu Gln Lys Leu
        435                 440                 445

Ser Gly Gln Glu Trp Arg Gly Leu Arg Asn Ile Phe Glu Ile Tyr Leu
    450                 455                 460

Ser Tyr Asn Lys Tyr Leu Gln Leu Ser Thr Ser Ser Phe Ala Leu Val
465                 470                 475                 480

Pro Ser Leu Gln Arg Leu Met Leu Arg Arg Val Ala Leu Lys Asn Val
                485                 490                 495

Asp Ile Ser Pro Ser Pro Phe Arg Pro Leu Arg Asn Leu Thr Ile Leu
            500                 505                 510

Asp Leu Ser Asn Asn Asn Ile Ala Asn Ile Asn Glu Asp Leu Leu Glu
        515                 520                 525

Gly Leu Glu Asn Leu Glu Ile Leu Asp Phe Gln His Asn Asn Leu Ala
    530                 535                 540

Arg Leu Trp Lys Arg Ala Asn Pro Gly Gly Pro Val Asn Phe Leu Lys
545                 550                 555                 560

Gly Leu Ser His Leu His Ile Leu Asn Leu Glu Ser Asn Gly Leu Asp
                565                 570                 575

Glu Ile Pro Val Gly Val Phe Lys Asn Leu Phe Glu Leu Lys Ser Ile
            580                 585                 590

Asn Leu Gly Leu Asn Asn Leu Asn Lys Leu Glu Pro Phe Ile Phe Asp
        595                 600                 605

Asp Gln Thr Ser Leu Arg Ser Leu Asn Leu Gln Lys Asn Leu Ile Thr
    610                 615                 620

Ser Val Glu Lys Asp Val Phe Gly Pro Pro Phe Gln Asn Leu Asn Ser
625                 630                 635                 640

Leu Asp Met Arg Phe Asn Pro Phe Asp Cys Thr Cys Glu Ser Ile Ser
                645                 650                 655

Trp Phe Val Asn Trp Ile Asn Gln Thr His Thr Asn Ile Phe Glu Leu
            660                 665                 670
```

```
Ser Thr His Tyr Leu Cys Asn Thr Pro His His Tyr Tyr Gly Phe Pro
        675                 680                 685

Leu Lys Leu Phe Asp Thr Ser Ser Cys Lys Asp Ser Ala Pro Phe Glu
    690                 695                 700

Leu Leu Phe Ile Ile Ser Thr Ser Met Leu Leu Val Phe Ile Leu Val
705                 710                 715                 720

Val Leu Leu Ile His Ile Glu Gly Trp Arg Ile Ser Phe Tyr Trp Asn
                725                 730                 735

Val Ser Val His Arg Ile Leu Gly Phe Lys Glu Ile Asp Thr Gln Ala
            740                 745                 750

Glu Gln Phe Glu Tyr Thr Ala Tyr Ile Ile His Ala His Lys Asp Arg
        755                 760                 765

Asp Trp Val Trp Glu His Phe Ser Pro Met Glu Glu Gln Asp Gln Ser
    770                 775                 780

Leu Lys Phe Cys Leu Glu Glu Arg Asp Phe Glu Ala Gly Val Leu Gly
785                 790                 795                 800

Leu Glu Ala Ile Val Asn Ser Ile Lys Arg Ser Arg Lys Ile Ile Phe
                805                 810                 815

Val Ile Thr His His Leu Leu Lys Asp Pro Leu Cys Arg Arg Phe Lys
            820                 825                 830

Val His His Ala Val Gln Gln Ala Ile Glu Gln Asn Leu Asp Ser Ile
        835                 840                 845

Ile Leu Ile Phe Leu Gln Asn Ile Pro Asp Tyr Lys Leu Asn His Ala
    850                 855                 860

Leu Cys Leu Arg Arg Gly Met Phe Lys Ser His Cys Ile Leu Asn Trp
865                 870                 875                 880

Pro Val Gln Lys Glu Arg Ile Asn Ala Phe His His Lys Leu Gln Val
                885                 890                 895

Ala Leu Gly Ser Arg Asn Ser Ala His
            900                 905
```

<210> 5
<211> 3352
<212> DNA
<213> Homo sapiens

<400> 5
aggctggtat aaaaatctta cttcctctat tctctgagcc gctgctgccc ctgtgggaag 60

ggacctcgag tgtgaagcat ccttccctgt agctgctgtc cagtctgccc gccagaccct 120

ctggagaagc ccctgccccc cagcatgggt ttctgccgca gcgccctgca cccgctgtct 180

ctcctggtgc aggccatcat gctggccatg accctggccc tgggtacctt gcctgccttc 240

ctaccctgtg agctccagcc ccacggcctg gtgaactgca actggctgtt cctgaagtct 300

gtgccccact tctccatggc agcaccccgt ggcaatgtca ccagcctttc cttgtcctcc 360

aaccgcatcc accacctcca tgattctgac tttgcccacc tgcccagcct gcggcatctc 420

aacctcaagt ggaactgccc gccggttggc ctcagcccca tgcacttccc ctgccacatg 480

accatcgagc ccagcacctt cttggctgtg cccaccctgg aagagctaaa cctgagctac 540

aacaacatca tgactgtgcc tgcgctgccc aaatccctca tatccctgtc cctcagccat 600

accaacatcc tgatgctaga ctctgccagc ctcgccggcc tgcatgccct gcgcttccta 660

ttcatggacg gcaactgtta ttacaagaac ccctgcaggc aggcactgga ggtggccccg 720

ggtgccctcc ttggcctggg caacctcacc cacctgtcac tcaagtacaa caacctcact 780

gtggtgcccc gcaacctgcc ttccagcctg gagtatctgc tgttgtccta caaccgcatc 840

gtcaaactgg cgcctgagga cctggccaat ctgaccgccc tgcgtgtgct cgatgtgggc 900

ggaaattgcc gccgctgcga ccacgctccc aacccctgca tggagtgccc tcgtcacttc 960

ccccagctac atcccgatac cttcagccac ctgagccgtc ttgaaggcct ggtgttgaag 1020

gacagttctc tctcctggct gaatgccagt tggttccgtg ggctgggaaa cctccgagtg 1080

ctggacctga gtgagaactt cctctacaaa tgcatcacta aaaccaaggc cttccagggc 1140

ctaacacagc tgcgcaagct taacctgtcc ttcaattacc aaaagagggt gtcctttgcc 1200

cacctgtctc tggccccttc cttcgggagc ctggtcgccc tgaaggagct ggacatgcac 1260

ggcatcttct tccgctcact cgatgagacc acgctccggc cactggcccg cctgcccatg 1320

ctccagactc tgcgtctgca gatgaacttc atcaaccagg cccagctcgg catcttcagg 1380

gccttccctg gcctgcgcta cgtggacctg tcggacaacc gcatcagcgg agcttcggag 1440

ctgacagcca ccatggggga ggcagatgga ggggagaagg tctggctgca gcctggggac 1500

cttgctccgg ccccagtgga cactcccagc tctgaagact tcaggcccaa ctgcagcacc 1560

ctcaacttca ccttggatct gtcacggaac aacctggtga ccgtgcagcc ggagatgttt 1620

gcccagctct cgcacctgca gtgcctgcgc ctgagccaca actgcatctc gcaggcagtc 1680

aatggctccc agttcctgcc gctgaccggt ctgcaggtgc tagacctgtc ccgcaataag 1740

ctggacctct accacgagca ctcattcacg gagctaccgc gactggaggc cctggacctc 1800

agctacaaca gccagccctt tggcatgcag ggcgtgggcc acaacttcag cttcgtggct 1860

cacctgcgca ccctgcgcca cctcagcctg gcccacaaca acatccacag ccaagtgtcc 1920

cagcagctct gcagtacgtc gctgcgggcc ctggacttca gcggcaatgc actgggccat 1980

atgtgggccg agggagacct ctatctgcac ttcttccaag gcctgagcgg tttgatctgg 2040

ctggacttgt cccagaaccg cctgcacacc ctcctgcccc aaaccctgcg caacctcccc 2100

aagagcctac aggtgctgcg tctccgtgac aattacctgg ccttctttaa gtggtggagc 2160

ctccacttcc tgcccaaact ggaagtcctc gacctggcag gaaaccggct gaaggccctg 2220

accaatggca gcctgcctgc tggcacccgg ctccggaggc tggatgtcag ctgcaacagc 2280

atcagcttcg tggcccccgg cttcttttcc aaggccaagg agctgcgaga gctcaacctt 2340

agcgccaacg ccctcaagac agtggaccac tcctggtttg ggcccctggc gagtgccctg 2400

caaatactag atgtaagcgc caaccctctg cactgcgcct gtggggcggc ctttatggac 2460

ttcctgctgg aggtgcaggc tgccgtgccc ggtctgccca gccgggtgaa gtgtggcagt 2520

ccgggccagc tccagggcct cagcatcttt gcacaggacc tgcgcctctg cctggatgag 2580

gccctctcct gggactgttt cgccctctcg ctgctggctg tggctctggg cctgggtgtg 2640

cccatgctgc atcacctctg tggctgggac ctctggtact gcttccacct gtgcctggcc 2700

tggcttccct ggcgggggcg gcaaagtggg cgagatgagg atgccctgcc ctacgatgcc 2760

ttcgtggtct tcgacaaaac gcagagcgca gtggcagact gggtgtacaa cgagcttcgg 2820

gggcagctgg aggagtgccg tgggcgctgg gcactccgcc tgtgcctgga ggaacgcgac 2880

tggctgcctg gcaaaaccct ctttgagaac ctgtgggcct cggtctatgg cagccgcaag 2940

acgctgtttg tgctggccca cacggaccgg gtcagtggtc tcttgcgcgc cagcttcctg 3000

ctggcccagc agcgcctgct ggaggaccgc aaggacgtcg tggtgctggt gatcctgagc 3060

cctgacggcc gccgctcccg ctacgtgcgg ctgcgccagc gcctctgccg ccagagtgtc 3120

ctcctctggc cccaccagcc cagtggtcag cgcagcttct gggcccagct gggcatggcc 3180

ctgaccaggg acaaccacca cttctataac cggaacttct gccagggacc cacggccgaa 3240

tagccgtgag ccggaatcct gcacggtgcc acctccacac tcacctcacc tctgcctgcc 3300

tggtctgacc ctcccctgct cgcctccctc accccacacc tgacacagag ca 3352

<210> 6
<211> 1032
<212> PRT
<213> Homo sapiens

<400> 6

Met Gly Phe Cys Arg Ser Ala Leu His Pro Leu Ser Leu Leu Val Gln
1 5 10 15

Ala Ile Met Leu Ala Met Thr Leu Ala Leu Gly Thr Leu Pro Ala Phe
20 25 30

Leu Pro Cys Glu Leu Gln Pro His Gly Leu Val Asn Cys Asn Trp Leu
35 40 45

Phe Leu Lys Ser Val Pro His Phe Ser Met Ala Ala Pro Arg Gly Asn
50 55 60

Val Thr Ser Leu Ser Leu Ser Ser Asn Arg Ile His His Leu His Asp
65 70 75 80

Ser Asp Phe Ala His Leu Pro Ser Leu Arg His Leu Asn Leu Lys Trp
85 90 95

Asn Cys Pro Pro Val Gly Leu Ser Pro Met His Phe Pro Cys His Met
100 105 110

Thr Ile Glu Pro Ser Thr Phe Leu Ala Val Pro Thr Leu Glu Glu Leu
115 120 125

Asn Leu Ser Tyr Asn Asn Ile Met Thr Val Pro Ala Leu Pro Lys Ser
130 135 140

Leu Ile Ser Leu Ser Leu Ser His Thr Asn Ile Leu Met Leu Asp Ser
145 150 155 160

Ala Ser Leu Ala Gly Leu His Ala Leu Arg Phe Leu Phe Met Asp Gly
165 170 175

Asn Cys Tyr Tyr Lys Asn Pro Cys Arg Gln Ala Leu Glu Val Ala Pro
180 185 190

Gly Ala Leu Leu Gly Leu Gly Asn Leu Thr His Leu Ser Leu Lys Tyr

```
        195                 200                 205

Asn Asn Leu Thr Val Val Pro Arg Asn Leu Pro Ser Ser Leu Glu Tyr
    210                 215                 220

Leu Leu Leu Ser Tyr Asn Arg Ile Val Lys Leu Ala Pro Glu Asp Leu
225                 230                 235                 240

Ala Asn Leu Thr Ala Leu Arg Val Leu Asp Val Gly Gly Asn Cys Arg
                245                 250                 255

Arg Cys Asp His Ala Pro Asn Pro Cys Met Glu Cys Pro Arg His Phe
            260                 265                 270

Pro Gln Leu His Pro Asp Thr Phe Ser His Leu Ser Arg Leu Glu Gly
        275                 280                 285

Leu Val Leu Lys Asp Ser Ser Leu Ser Trp Leu Asn Ala Ser Trp Phe
    290                 295                 300

Arg Gly Leu Gly Asn Leu Arg Val Leu Asp Leu Ser Glu Asn Phe Leu
305                 310                 315                 320

Tyr Lys Cys Ile Thr Lys Thr Lys Ala Phe Gln Gly Leu Thr Gln Leu
                325                 330                 335

Arg Lys Leu Asn Leu Ser Phe Asn Tyr Gln Lys Arg Val Ser Phe Ala
            340                 345                 350

His Leu Ser Leu Ala Pro Ser Phe Gly Ser Leu Val Ala Leu Lys Glu
        355                 360                 365

Leu Asp Met His Gly Ile Phe Phe Arg Ser Leu Asp Glu Thr Thr Leu
    370                 375                 380

Arg Pro Leu Ala Arg Leu Pro Met Leu Gln Thr Leu Arg Leu Gln Met
385                 390                 395                 400

Asn Phe Ile Asn Gln Ala Gln Leu Gly Ile Phe Arg Ala Phe Pro Gly
                405                 410                 415

Leu Arg Tyr Val Asp Leu Ser Asp Asn Arg Ile Ser Gly Ala Ser Glu
            420                 425                 430
```

```
Leu Thr Ala Thr Met Gly Glu Ala Asp Gly Gly Glu Lys Val Trp Leu
        435                 440                 445

Gln Pro Gly Asp Leu Ala Pro Ala Pro Val Asp Thr Pro Ser Ser Glu
    450                 455                 460

Asp Phe Arg Pro Asn Cys Ser Thr Leu Asn Phe Thr Leu Asp Leu Ser
465                 470                 475                 480

Arg Asn Asn Leu Val Thr Val Gln Pro Glu Met Phe Ala Gln Leu Ser
                485                 490                 495

His Leu Gln Cys Leu Arg Leu Ser His Asn Cys Ile Ser Gln Ala Val
            500                 505                 510

Asn Gly Ser Gln Phe Leu Pro Leu Thr Gly Leu Gln Val Leu Asp Leu
        515                 520                 525

Ser Arg Asn Lys Leu Asp Leu Tyr His Glu His Ser Phe Thr Glu Leu
    530                 535                 540

Pro Arg Leu Glu Ala Leu Asp Leu Ser Tyr Asn Ser Gln Pro Phe Gly
545                 550                 555                 560

Met Gln Gly Val Gly His Asn Phe Ser Phe Val Ala His Leu Arg Thr
                565                 570                 575

Leu Arg His Leu Ser Leu Ala His Asn Asn Ile His Ser Gln Val Ser
            580                 585                 590

Gln Gln Leu Cys Ser Thr Ser Leu Arg Ala Leu Asp Phe Ser Gly Asn
        595                 600                 605

Ala Leu Gly His Met Trp Ala Glu Gly Asp Leu Tyr Leu His Phe Phe
    610                 615                 620

Gln Gly Leu Ser Gly Leu Ile Trp Leu Asp Leu Ser Gln Asn Arg Leu
625                 630                 635                 640

His Thr Leu Leu Pro Gln Thr Leu Arg Asn Leu Pro Lys Ser Leu Gln
                645                 650                 655

Val Leu Arg Leu Arg Asp Asn Tyr Leu Ala Phe Phe Lys Trp Trp Ser
            660                 665                 670
```

```
Leu His Phe Leu Pro Lys Leu Glu Val Leu Asp Leu Ala Gly Asn Arg
        675                 680                 685

Leu Lys Ala Leu Thr Asn Gly Ser Leu Pro Ala Gly Thr Arg Leu Arg
    690                 695                 700

Arg Leu Asp Val Ser Cys Asn Ser Ile Ser Phe Val Ala Pro Gly Phe
705                 710                 715                 720

Phe Ser Lys Ala Lys Glu Leu Arg Glu Leu Asn Leu Ser Ala Asn Ala
                725                 730                 735

Leu Lys Thr Val Asp His Ser Trp Phe Gly Pro Leu Ala Ser Ala Leu
            740                 745                 750

Gln Ile Leu Asp Val Ser Ala Asn Pro Leu His Cys Ala Cys Gly Ala
        755                 760                 765

Ala Phe Met Asp Phe Leu Leu Glu Val Gln Ala Ala Val Pro Gly Leu
    770                 775                 780

Pro Ser Arg Val Lys Cys Gly Ser Pro Gly Gln Leu Gln Gly Leu Ser
785                 790                 795                 800

Ile Phe Ala Gln Asp Leu Arg Leu Cys Leu Asp Glu Ala Leu Ser Trp
                805                 810                 815

Asp Cys Phe Ala Leu Ser Leu Leu Ala Val Ala Leu Gly Leu Gly Val
            820                 825                 830

Pro Met Leu His His Leu Cys Gly Trp Asp Leu Trp Tyr Cys Phe His
        835                 840                 845

Leu Cys Leu Ala Trp Leu Pro Trp Arg Gly Arg Gln Ser Gly Arg Asp
    850                 855                 860

Glu Asp Ala Leu Pro Tyr Asp Ala Phe Val Val Phe Asp Lys Thr Gln
865                 870                 875                 880

Ser Ala Val Ala Asp Trp Val Tyr Asn Glu Leu Arg Gly Gln Leu Glu
                885                 890                 895

Glu Cys Arg Gly Arg Trp Ala Leu Arg Leu Cys Leu Glu Glu Arg Asp
            900                 905                 910
```

```
Trp Leu Pro Gly Lys Thr Leu Phe Glu Asn Leu Trp Ala Ser Val Tyr
        915                 920                 925

Gly Ser Arg Lys Thr Leu Phe Val Leu Ala His Thr Asp Arg Val Ser
    930                 935                 940

Gly Leu Leu Arg Ala Ser Phe Leu Leu Ala Gln Gln Arg Leu Leu Glu
945                 950                 955                 960

Asp Arg Lys Asp Val Val Val Leu Val Ile Leu Ser Pro Asp Gly Arg
                965                 970                 975

Arg Ser Arg Tyr Val Arg Leu Arg Gln Arg Leu Cys Arg Gln Ser Val
            980                 985                 990

Leu Leu Trp Pro His Gln Pro Ser  Gly Gln Arg Ser Phe  Trp Ala Gln
        995                 1000                1005

Leu Gly  Met Ala Leu Thr Arg  Asp Asn His His Phe  Tyr Asn Arg
    1010                1015                1020

Asn Phe  Cys Gln Gly Pro Thr  Ala Glu
    1025                1030
```

<210> 7
<211> 3200
<212> DNA
<213> Mus musculus

<400> 7

```
tgtcagaggg agcctcggga gaatcctcca tctcccaaca tggttctccg tcgaaggact     60

ctgcacccct tgtccctcct ggtacaggct gcagtgctgg ctgagactct ggccctgggt    120

accctgcctg ccttcctacc ctgtgagctg aagcctcatg gcctggtgga ctgcaattgg    180

ctgttcctga agtctgtacc ccgtttctct gcggcagcat cctgctccaa catcacccgc    240

ctctccttga tctccaaccg tatccaccac ctgcacaact ccgacttcgt ccacctgtcc    300

aacctgcggc agctgaacct caagtggaac tgtccaccca ctggccttag ccccctgcac    360

ttctcttgcc acatgaccat tgagcccaga accttcctgg ctatgcgtac actggaggag    420

ctgaacctga gctataatgg tatcaccact gtgccccgac tgcccagctc cctggtgaat    480

ctgagcctga gccacaccaa catcctggtt ctagatgcta acagcctcgc cggcctatac    540

agcctgcgcg ttctcttcat ggacgggaac tgctactaca agaacccctg cacaggagcg    600
```

gtgaaggtga ccccaggcgc cctcctgggc ctgagcaatc tcacccatct gtctctgaag 660

tataacaacc tcacaaaggt gccccgccaa ctgcccccca gcctggagta cctcctggtg 720

tcctataacc tcattgtcaa gctggggcct gaagacctgg ccaatctgac ctcccttcga 780

gtacttgatg tgggtgggaa ttgccgtcgc tgcgaccatg cccccaatcc ctgtatagaa 840

tgtggccaaa agtccctcca cctgcaccct gagaccttcc atcacctgag ccatctggaa 900

ggcctggtgc tgaaggacag ctctctccat acactgaact cttcctggtt ccaaggtctg 960

gtcaacctct cggtgctgga cctaagcgag aactttctct atgaaagcat caaccacacc 1020

aatgcctttc agaacctaac ccgcctgcgc aagctcaacc tgtccttcaa ttaccgcaag 1080

aaggtatcct ttgcccgcct ccacctggca agttccttca agaacctggt gtcactgcag 1140

gagctgaaca tgaacggcat cttcttccgc tcgctcaaca agtacacgct cagatggctg 1200

gccgatctgc ccaaactcca cactctgcat cttcaaatga acttcatcaa ccaggcacag 1260

ctcagcatct ttggtacctt ccgagccctt cgctttgtgg acttgtcaga caatcgcatc 1320

agtgggcctt caacgctgtc agaagccacc cctgaagagg cagatgatgc agagcaggag 1380

gagctgttgt ctgcggatcc tcacccagct ccactgagca cccctgcttc taagaacttc 1440

atggacaggt gtaagaactt caagttcacc atggacctgt ctcggaacaa cctggtgact 1500

atcaagccag agatgtttgt caatctctca cgcctccagt gtcttagcct gagccacaac 1560

tccattgcac aggctgtcaa tggctctcag ttcctgccgc tgactaatct gcaggtgctg 1620

gacctgtccc ataacaaact ggacttgtac cactggaaat cgttcagtga gctaccacag 1680

ttgcaggccc tggacctgag ctacaacagc cagcccttta gcatgaaggg tataggccac 1740

aatttcagtt ttgtggccca tctgtccatg ctacacagcc ttagcctggc acacaatgac 1800

attcataccc gtgtgtcctc acatctcaac agcaactcag tgaggtttct tgacttcagc 1860

ggcaacggta tgggccgcat gtgggatgag gggggccttt atctccattt cttccaaggc 1920

ctgagtggcc tgctgaagct ggacctgtct caaaataacc tgcatatcct ccggccccag 1980

aaccttgaca acctccccaa gagcctgaag ctgctgagcc tccgagacaa ctacctatct 2040

ttctttaact ggaccagtct gtccttcctg cccaacctgg aagtcctaga cctggcaggc 2100

aaccagctaa aggccctgac caatggcacc ctgcctaatg gcaccctcct ccagaaactg 2160

gatgtcagca gcaacagtat cgtctctgtg gtcccagcct tcttcgctct ggcggtcgag 2220

ctgaaagagg tcaacctcag ccacaacatt ctcaagacgg tggatcgctc ctggtttggg 2280

cccattgtga tgaacctgac agttctagac gtgagaagca accctctgca ctgtgcctgt 2340

ggggcagcct tcgtagactt actgttggag gtgcagacca aggtgcctgg cctggctaat 2400

ggtgtgaagt gtggcagccc cggccagctg cagggccgta gcatcttcgc acaggacctg 2460

cggctgtgcc tggatgaggt cctctcttgg gactgctttg gcctttcact cttggctgtg 2520

gccgtgggca tggtggtgcc tatactgcac catctctgcg gctgggacgt ctggtactgt 2580

tttcatctgt gcctggcatg gctacctttg ctggcccgca gccgacgcag cgcccaagct 2640

ctcccctatg atgccttcgt ggtgttcgat aaggcacaga gcgcagttgc ggactgggtg 2700

tataacgagc tgcgggtgcg gctggaggag cggcgcggtc gccgagccct acgcttgtgt 2760

ctggaggacc gagattggct gcctggccag acgctcttcg agaacctctg ggcttccatc 2820

tatgggagcc gcaagactct atttgtgctg gcccacacgg accgcgtcag tggcctcctg 2880

cgcaccagct tcctgctggc tcagcagcgc ctgttggaag accgcaagga cgtggtggtg 2940

ttggtgatcc tgcgtccgga tgcccaccgc tcccgctatg tgcgactgcg ccagcgtctc 3000

tgccgccaga gtgtgctctt ctggccccag cagcccaacg ggcaggggggg cttctgggcc 3060

cagctgagta cagccctgac tagggacaac cgccacttct ataaccagaa cttctgccgg 3120

ggacctacag cagaatagct cagagcaaca gctggaaaca gctgcatctt catgcctggt 3180

tcccgagttg ctctgcctgc 3200

<210> 8
<211> 1032
<212> PRT
<213> Mus musculus

<400> 8

Met Val Leu Arg Arg Arg Thr Leu His Pro Leu Ser Leu Leu Val Gln
1 5 10 15

Ala Ala Val Leu Ala Glu Thr Leu Ala Leu Gly Thr Leu Pro Ala Phe
20 25 30

Leu Pro Cys Glu Leu Lys Pro His Gly Leu Val Asp Cys Asn Trp Leu
35 40 45

Phe Leu Lys Ser Val Pro Arg Phe Ser Ala Ala Ala Ser Cys Ser Asn
50 55 60

Ile Thr Arg Leu Ser Leu Ile Ser Asn Arg Ile His His Leu His Asn
65 70 75 80

Ser Asp Phe Val His Leu Ser Asn Leu Arg Gln Leu Asn Leu Lys Trp

```
                85                  90                  95

Asn Cys Pro Pro Thr Gly Leu Ser Pro Leu His Phe Ser Cys His Met
            100                 105                 110

Thr Ile Glu Pro Arg Thr Phe Leu Ala Met Arg Thr Leu Glu Glu Leu
        115                 120                 125

Asn Leu Ser Tyr Asn Gly Ile Thr Thr Val Pro Arg Leu Pro Ser Ser
    130                 135                 140

Leu Val Asn Leu Ser Leu Ser His Thr Asn Ile Leu Val Leu Asp Ala
145                 150                 155                 160

Asn Ser Leu Ala Gly Leu Tyr Ser Leu Arg Val Leu Phe Met Asp Gly
                165                 170                 175

Asn Cys Tyr Tyr Lys Asn Pro Cys Thr Gly Ala Val Lys Val Thr Pro
            180                 185                 190

Gly Ala Leu Leu Gly Leu Ser Asn Leu Thr His Leu Ser Leu Lys Tyr
        195                 200                 205

Asn Asn Leu Thr Lys Val Pro Arg Gln Leu Pro Pro Ser Leu Glu Tyr
    210                 215                 220

Leu Leu Val Ser Tyr Asn Leu Ile Val Lys Leu Gly Pro Glu Asp Leu
225                 230                 235                 240

Ala Asn Leu Thr Ser Leu Arg Val Leu Asp Val Gly Gly Asn Cys Arg
                245                 250                 255

Arg Cys Asp His Ala Pro Asn Pro Cys Ile Glu Cys Gly Gln Lys Ser
            260                 265                 270

Leu His Leu His Pro Glu Thr Phe His His Leu Ser His Leu Glu Gly
        275                 280                 285

Leu Val Leu Lys Asp Ser Ser Leu His Thr Leu Asn Ser Ser Trp Phe
    290                 295                 300

Gln Gly Leu Val Asn Leu Ser Val Leu Asp Leu Ser Glu Asn Phe Leu
305                 310                 315                 320
```

```
Tyr Glu Ser Ile Asn His Thr Asn Ala Phe Gln Asn Leu Thr Arg Leu
                325                 330                 335

Arg Lys Leu Asn Leu Ser Phe Asn Tyr Arg Lys Lys Val Ser Phe Ala
            340                 345                 350

Arg Leu His Leu Ala Ser Ser Phe Lys Asn Leu Val Ser Leu Gln Glu
        355                 360                 365

Leu Asn Met Asn Gly Ile Phe Phe Arg Ser Leu Asn Lys Tyr Thr Leu
    370                 375                 380

Arg Trp Leu Ala Asp Leu Pro Lys Leu His Thr Leu His Leu Gln Met
385                 390                 395                 400

Asn Phe Ile Asn Gln Ala Gln Leu Ser Ile Phe Gly Thr Phe Arg Ala
                405                 410                 415

Leu Arg Phe Val Asp Leu Ser Asp Asn Arg Ile Ser Gly Pro Ser Thr
            420                 425                 430

Leu Ser Glu Ala Thr Pro Glu Glu Ala Asp Asp Ala Glu Gln Glu Glu
        435                 440                 445

Leu Leu Ser Ala Asp Pro His Pro Ala Pro Leu Ser Thr Pro Ala Ser
    450                 455                 460

Lys Asn Phe Met Asp Arg Cys Lys Asn Phe Lys Phe Thr Met Asp Leu
465                 470                 475                 480

Ser Arg Asn Asn Leu Val Thr Ile Lys Pro Glu Met Phe Val Asn Leu
                485                 490                 495

Ser Arg Leu Gln Cys Leu Ser Leu Ser His Asn Ser Ile Ala Gln Ala
            500                 505                 510

Val Asn Gly Ser Gln Phe Leu Pro Leu Thr Asn Leu Gln Val Leu Asp
        515                 520                 525

Leu Ser His Asn Lys Leu Asp Leu Tyr His Trp Lys Ser Phe Ser Glu
    530                 535                 540

Leu Pro Gln Leu Gln Ala Leu Asp Leu Ser Tyr Asn Ser Gln Pro Phe
545                 550                 555                 560
```

```
Ser Met Lys Gly Ile Gly His Asn Phe Ser Phe Val Ala His Leu Ser
                565                 570                 575

Met Leu His Ser Leu Ser Leu Ala His Asn Asp Ile His Thr Arg Val
            580                 585                 590

Ser Ser His Leu Asn Ser Asn Ser Val Arg Phe Leu Asp Phe Ser Gly
        595                 600                 605

Asn Gly Met Gly Arg Met Trp Asp Glu Gly Gly Leu Tyr Leu His Phe
    610                 615                 620

Phe Gln Gly Leu Ser Gly Leu Leu Lys Leu Asp Leu Ser Gln Asn Asn
625                 630                 635                 640

Leu His Ile Leu Arg Pro Gln Asn Leu Asp Asn Leu Pro Lys Ser Leu
                645                 650                 655

Lys Leu Leu Ser Leu Arg Asp Asn Tyr Leu Ser Phe Phe Asn Trp Thr
            660                 665                 670

Ser Leu Ser Phe Leu Pro Asn Leu Glu Val Leu Asp Leu Ala Gly Asn
        675                 680                 685

Gln Leu Lys Ala Leu Thr Asn Gly Thr Leu Pro Asn Gly Thr Leu Leu
    690                 695                 700

Gln Lys Leu Asp Val Ser Ser Asn Ser Ile Val Ser Val Val Pro Ala
705                 710                 715                 720

Phe Phe Ala Leu Ala Val Glu Leu Lys Glu Val Asn Leu Ser His Asn
                725                 730                 735

Ile Leu Lys Thr Val Asp Arg Ser Trp Phe Gly Pro Ile Val Met Asn
            740                 745                 750

Leu Thr Val Leu Asp Val Arg Ser Asn Pro Leu His Cys Ala Cys Gly
        755                 760                 765

Ala Ala Phe Val Asp Leu Leu Leu Glu Val Gln Thr Lys Val Pro Gly
    770                 775                 780

Leu Ala Asn Gly Val Lys Cys Gly Ser Pro Gly Gln Leu Gln Gly Arg
785                 790                 795                 800
```

```
Ser Ile Phe Ala Gln Asp Leu Arg Leu Cys Leu Asp Glu Val Leu Ser
                805                 810                 815

Trp Asp Cys Phe Gly Leu Ser Leu Leu Ala Val Ala Val Gly Met Val
            820                 825                 830

Val Pro Ile Leu His His Leu Cys Gly Trp Asp Val Trp Tyr Cys Phe
        835                 840                 845

His Leu Cys Leu Ala Trp Leu Pro Leu Leu Ala Arg Ser Arg Arg Ser
    850                 855                 860

Ala Gln Ala Leu Pro Tyr Asp Ala Phe Val Val Phe Asp Lys Ala Gln
865                 870                 875                 880

Ser Ala Val Ala Asp Trp Val Tyr Asn Glu Leu Arg Val Arg Leu Glu
                885                 890                 895

Glu Arg Arg Gly Arg Arg Ala Leu Arg Leu Cys Leu Glu Asp Arg Asp
            900                 905                 910

Trp Leu Pro Gly Gln Thr Leu Phe Glu Asn Leu Trp Ala Ser Ile Tyr
        915                 920                 925

Gly Ser Arg Lys Thr Leu Phe Val Leu Ala His Thr Asp Arg Val Ser
    930                 935                 940

Gly Leu Leu Arg Thr Ser Phe Leu Leu Ala Gln Gln Arg Leu Leu Glu
945                 950                 955                 960

Asp Arg Lys Asp Val Val Val Leu Val Ile Leu Arg Pro Asp Ala His
                965                 970                 975

Arg Ser Arg Tyr Val Arg Leu Arg Gln Arg Leu Cys Arg Gln Ser Val
            980                 985                 990

Leu Phe Trp Pro Gln Gln Pro Asn  Gly Gln Gly Gly Phe  Trp Ala Gln
        995                 1000                1005

Leu Ser  Thr Ala Leu Thr Arg  Asp Asn Arg His Phe  Tyr Asn Gln
    1010                1015                1020

Asn Phe  Cys Arg Gly Pro Thr  Ala Glu
```

1025 1030

<210> 9
<211> 42
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 9
gaaactcgag ccaccatgag acagactttg ccttgtatct ac 42

<210> 10
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 10
gaaagaattc ttaatgtaca gagtttttgg atccaag 37

<210> 11
<211> 670
<212> DNA
<213> Homo sapiens

<400> 11
agaaaaattt taaaaaatta ttcattcata tttttaggag ttttgaatga ttggatatgt 60

aattatattc atattattaa tgtgtatcta tatagatttt tattttgcat atgtactttg 120

atacaaaatt tacatgaaca aattacacta aaagttattc cacaaatata cttatcaaat 180

taagttaaat gtcaatagct tttaaactta aattttagtt taacttttct gtcattcttt 240

actttgaata aaaagagcaa actttgtagt ttttatctgt gaagtagagg tatacgtaat 300

atacataaat agatatgcca aatctgtgtt attaaaattt catgaagatt tcaattagaa 360

aaaaatacca taaaaggctt tgagtgcagg tgaaaaatag gcaatgatga aaaaaaatga 420

aaaacttttt aaacacatgt agagagtgcg taaagaaagc aaaaacagag atagaaagta 480

caactaggga atttagaaaa tggaaattag tatgttcact atttaagacc tatgcacaga 540

gcaaagtctt cagaaaacct agaggccgaa gttcaaggtt atccatctca agtagcctag 600

caatatttgc aacatcccaa tggccctgtc cttttcttta ctgatggccg tgctggtgct 660

cagctacaaa 670

<210> 12

<211> 300
<212> DNA
<213> Homo sapiens

<400> 12
ttctcaggtc gtttgctttc ctttgctttc tcccaagtct tgttttacaa tttgctttag 60

tcattcactg aaactttaaa aaacattaga aaacctcaca gtttgtaaat ctttttccct 120

attatatata tcataagata ggagcttaaa taaagagttt tagaaactac taaaatgtaa 180

atgacatagg aaaactgaaa gggagaagtg aaagtgggaa attcctctga atagagagag 240

gaccatctca tataaatagg ccatacccac ggagaaagga cattctaact gcaacctttc 300

<210> 13
<211> 1031
<212> DNA
<213> Homo sapiens

<400> 13
agaaggcctt acagtgagat gggatcccag tatttattga gtttcctcat tcataaaatg 60

gggataataa tagtaaatga gttgacacgc gctaagacag tggaatagtg gctggcacag 120

ataagccctc ggtaaatggt agccaataat gatagagtat gctgtaagat atctttctct 180

ccctctgctt ctcaacaagt ctctaatcaa ttattccact ttataaacaa ggaaatagaa 240

ctcaaagaca ttaagcactt ttcccaaagg tcgcttagca agtaaatggg agagacccta 300

tgaccaggat gaaagcaaga aattcccaca agaggactca ttccaactca tatcttgtga 360

aaaggttccc aatgcccagc tcagatcaac tgcctcaatt tacagtgtga gtgtgctcac 420

ctcctttggg gactgtatat ccagaggacc ctcctcaata aaacacttta taaataacat 480

ccttccatgg atgagggaaa ggaggtaaga tctgtaatga ataagcagga actttgaaga 540

ctcagtgact cagtgagtaa taaagactca gtgacttctg atcctgtcct aactgccact 600

ccttgttgtc cccaagaaag cggcttcctg ctctctgagg aggacccctt ccctggaagg 660

taaaactaag gatgtcagca gagaaatttt tccaccattg gtgcttggtc aaagaggaaa 720

ctgatgagct cactctagat gagagagcag tgagggagag acagagactc gaatttccgg 780

aggctatttc agttttcttt tccgttttgt gcaatttcac ttatgatacc ggccaatgct 840

tggttgctat tttggaaact ccccttaggg gatgcccctc aactggccct ataaagggcc 900

agcctgagct gcagaggatt cctgcagagg atcaagacag cacgtggacc tcgcacagcc 960

tctcccacag gtaccatgaa ggtctccgcg gcagccctcg ctgtcatcct cattgctact 1020

gccctctgcg c 1031

<210> 14
<211> 401
<212> DNA
<213> Homo sapiens

<400> 14
gatctgtaat gaataagcag gaactttgaa gactcagtga ctcagtgagt aataaagact 60

cagtgacttc tgatcctgtc ctaactgcca ctccttgttg tcccaagaaa gcggcttcct 120

gctctctgag gaggacccct tccctggaag gtaaaactaa ggatgtcagc agagaaattt 180

ttccaccatt ggtgcttggt caaagaggaa actgatgagc tcactctaga tgagagagca 240

gtgagggaga gacagagact cgaatttccg gagctatttc agttttcttt tccgttttgt 300

gcaatttcac ttatgatacc ggccaatgct tggttgctat tttggaaact ccccttaggg 360

gatgcccctc aactggccct ataaagggcc agcctgagct g 401

<210> 15
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 15
tcgtcgtttt gtcgttttgt cgtt 24

<210> 16
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 16
tgctgctttt gtgcttttgt gctt 24

<210> 17
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<220>
<221> misc_feature
<222> (2)..(2)
<223> n = m5c

<220>
<221> misc_feature
<222> (5)..(5)
<223> n = m5c

<220>
<221> misc_feature
<222> (13)..(13)
<223> n = m5c

<220>
<221> misc_feature
<222> (21)..(21)
<223> n = m5c

<400> 17
tngtngtttt gtngttttgt ngtt 24

<210> 18
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 18
tccatgacgt tcctgatgct 20

<210> 19
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 19
tccatgagct tcctgatgct 20

<210> 20
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<220>
<221> misc_feature
<222> (8)..(8)
<223> n = m5c

<400> 20
tccatgangt tcctgatgct 20

<210> 21
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 21
gcgactggct gcatggcaaa accctctttg 30

<210> 22
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 22
caaagagggt tttgccatgc agccagtcgc 30

<210> 23
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 23
cgagattggc tgcatggcca gacgctcttc 30

<210> 24
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 24
gaagagcgtc tggccatgca gccaatctcg 30

<210> 25
<211> 15
<212> DNA
<213> Artificial sequence

<220>

<223> Synthetic oligonucleotide

<400> 25
ggcctcagca tcttt 15

<210> 26
<211> 15
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 26
ggcctatcga ttttt 15

<210> 27
<211> 15
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 27
gggttcccag tgaga 15

<210> 28
<211> 15
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 28
gggttatcga ttaga 15

<210> 29
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 29
cagctccagg gcctatcgat ttttgcacag gacc 34

<210> 30
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 30
ggtcctgtgc aaaaatcgat aggccctgga gctg 34

<210> 31
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 31
tccatgacgt ttttgatgtt 20

<210> 32
<211> 18
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 32
tccatgacgt ttttgatg 18

<210> 33
<211> 16
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 33
tccatgacgt ttttga 16

<210> 34
<211> 14
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 34
tccatgacgt tttt 14

<210> 35
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 35
tccatgacgt ttttgatgtt 20

<210> 36
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 36
tcgtcgtttt gtcgttttgt cgtt 24

<210> 37
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 37
tcgtcgtttt gtcgttttgt cgtt 24

<210> 38
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 38
tccatgacgt ttttgatgtt 20

<210> 39
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 39
aagcgaaaat gaaattgact 20

<210> 40
<211> 24
<212> DNA

<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 40
accatggacg aactgtttcc cctc 24

<210> 41
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 41
accatggacg acctgtttcc cctc 24

<210> 42
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 42
accatggacg agctgtttcc cctc 24

<210> 43
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 43
accatggacg atctgtttcc cctc 24

<210> 44
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 44
accatggacg gtctgtttcc cctc 24

<210> 45
<211> 24

<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 45
accatggacg tactgtttcc cctc 24

<210> 46
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 46
accatggacg ttctgtttcc cctc 24

<210> 47
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 47
agcgggggcg agcgggggcg 20

<210> 48
<211> 18
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 48
agctatgacg ttccaagg 18

<210> 49
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 49
atcgactctc gagcgttctc 20

<210> 50

<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 50
atgacgttcc tgacgtt 17

<210> 51
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 51
atggaaggtc caacgttctc 20

<210> 52
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 52
atggaaggtc cagcgttctc 20

<210> 53
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 53
atggactctc cagcgttctc 20

<210> 54
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 54
atggaggctc catcgttctc 20

<210> 55
<211> 15
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 55
cacgttgagg ggcat 15

<210> 56
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 56
caggcataac ggttccgtag 20

<210> 57
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 57
ctgatttccc cgaaatgatg 20

<210> 58
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 58
gagaacgatg gaccttccat 20

<210> 59
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 59
gagaacgctc cagcactgat 20

<210> 60
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 60
gagaacgctc gaccttccat 20

<210> 61
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 61
gagaacgctc gaccttcgat 20

<210> 62
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 62
gagaacgctg gaccttccat 20

<210> 63
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 63
gattgcctga cgtcagagag 20

<210> 64
<211> 15
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 64
gcatgacgtt gagct 15

<210> 65
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 65
gcggcgggcg gcgcgcgccc 20

<210> 66
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 66
gcgtgcgttg tcgttgtcgt t 21

<210> 67
<211> 15
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 67
gctagacgtt agcgt 15

<210> 68
<211> 15
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 68
gctagacgtt agtgt 15

<210> 69
<211> 15
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 69

gctagatgtt agcgt 15

<210> 70
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 70
gcttgatgac tcagccggaa 20

<210> 71
<211> 18
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 71
ggaatgacgt tccctgtg 18

<210> 72
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 72
ggggtcaacg ttgacgggg 19

<210> 73
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 73
ggggtcagtc ttgacgggg 19

<210> 74
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 74
gtccatttcc cgtaaatctt 20

<210> 75
<211> 18
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 75
taccgcgtgc gaccctct 18

<210> 76
<211> 12
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 76
tcagcgtgcg cc 12

<210> 77
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 77
tccacgacgt tttcgacgtt 20

<210> 78
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 78
tccataacgt tcctgatgct 20

<210> 79
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 79
tccatagcgt tcctagcgtt 20

<210> 80
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 80
tccatcacgt gcctgatgct 20

<210> 81
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 81
tccatgacgg tcctgatgct 20

<210> 82
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 82
tccatgacgt ccctgatgct 20

<210> 83
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 83
tccatgacgt gcctgatgct 20

<210> 84
<211> 20
<212> DNA
<213> Artificial sequence

<220>

<223> Synthetic oligonucleotide

<400> 84
tccatgacgt tcctgacgtt 20

<210> 85
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 85
tccatgccgg tcctgatgct 20

<210> 86
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 86
tccatgcgtg cgtgcgtttt 20

<210> 87
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 87
tccatgcgtt gcgttgcgtt 20

<210> 88
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 88
tccatggcgg tcctgatgct 20

<210> 89
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 89
tccatgtcga tcctgatgct 20

<210> 90
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 90
tccatgtcgc tcctgatgct 20

<210> 91
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 91
tccatgtcgg tcctgatgct 20

<210> 92
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 92
tccatgtcgg tcctgctgat 20

<210> 93
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 93
tccatgtcgt ccctgatgct 20

<210> 94
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 94
tccatgtcgt tcctgatgct 20

<210> 95
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 95
tccatgtcgt tcctgtcgtt 20

<210> 96
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 96
tccatgtcgt ttttgtcgtt 20

<210> 97
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 97
tcctgacgtt cctgacgtt 19

<210> 98
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 98
tcctgtcgtt cctgtcgtt 19

<210> 99
<211> 20
<212> DNA

<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 99
tcctgtcgtt ccttgtcgtt 20

<210> 100
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 100
tcctgtcgtt ttttgtcgtt 20

<210> 101
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 101
tccttgtcgt tcctgtcgtt 20

<210> 102
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 102
tcgatcgggg cggggcgagc 20

<210> 103
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 103
tcgtcgctgt ctccgcttct t 21

<210> 104
<211> 27

<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 104
tcgtcgctgt ctccgcttct tcttgcc 27

<210> 105
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 105
tcgtcgctgt ctgcccttct t 21

<210> 106
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 106
tcgtcgctgt tgtcgtttct t 21

<210> 107
<211> 14
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 107
tcgtcgtcgt cgtt 14

<210> 108
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 108
tcgtcgttgt cgttgtcgtt 20

<210> 109

<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 109
tcgtcgttgt cgttttgtcg tt 22

<210> 110
<211> 18
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 110
tctcccagcg cgcgccat 18

<210> 111
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 111
tctcccagcg ggcgcat 17

<210> 112
<211> 18
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 112
tctcccagcg tgcgccat 18

<210> 113
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 113
tgcagattgc gcaatctgca 20

<210> 114
<211> 13
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 114
tgtcgttgtc gtt 13

<210> 115
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 115
tgtcgttgtc gttgtcgtt 19

<210> 116
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 116
tgtcgttgtc gttgtcgttg tcgtt 25

<210> 117
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide

<400> 117
tgtcgtttgt cgtttgtcgt t 21